

Alexander Balanescu

**Ein Beitrag zum Problem der
Selektion wichtiger Symptome im Zusammenhang
mit der computerunterstützten Diagnosefindung**

Dissertation
zur Erlangung des Grades eines Doktors
der Ingenieurwissenschaften
der Universität Bremen

Bremen, Mai 1978

Alexander Balanescu

**Ein Beitrag zum Problem der
Selektion wichtiger Symptome im Zusammenhang
mit der computerunterstützten Diagnosefindung**

Dissertation
zur Erlangung des Grades eines Doktors
der Ingenieurwissenschaften
der Universität Bremen

Diese Arbeit hat dem Promotionsausschuß Dr. ing. der Universität Bremen als Dissertation vorgelegen. An dem Promotionsverfahren haben Dobrivoje Popovic und Siegfried Boseck als Gutachter mitgewirkt. Das Kolloquium fand am 15. 8. 1977 statt.

Inhaltsverzeichnis

	Seite
0. Übersicht	1
1. Einleitung	4
2. Chronologische Entwicklung der mathematischen Methoden zur Diagnoseunterstützung	10
3. Auflistung der Fragestellungen	15
4. Klassifikationsverfahren zur Diagnoseunterstützung	21
4.1. Die Problemstellung	21
4.2. Bayes-Lösung	22
4.3. Minimax-Lösung	24
4.4. Aposteriori Bayes-Wahrscheinlichkeiten	24
4.5. Klassifikation in normalverteilten Populationen	26
5. Fehler beim Abschätzen und Klassifizieren mit vielen Variablen	29
5.1. Die Fehlklassifikationsrate als Funktion der Anzahl der Dimensionen	31
5.2. Systematische Fehler und redundante Variablen	36



	Seite
6. Bekannte Verfahren zur Reduktion der Dimensionen	40
6.1. Hauptkomponentenmethode	40
6.2. Methode des mehrfachen Korrelationskoeffizienten	42
6.3. Wilks'- Lambda - Test	43
6.4. Reduktion des Diskriminanzraumes	45
6.5. Sequentielle Kovarianzanalyse	46
7. Informationstheoretische Diskriminanzanalyse	48
7.1. Das Konzept der "Trenninformation" und der "informationstheoretischen Divergenz"	48
7.2. Informationstheoretische Erklärung der Hauptkomponentenmethode	50
7.3. Die informationstheoretische Divergenz als Kriterium für die Dimensionenreduktion	53
8. Untersuchungen mit Hilfe von Monte-Carlo-Experimenten	59
8.1. Aufbau des Symulationssystems und Beschreibung des Datenmaterials	60
8.2. Darstellung und Diskussion der Ergebnisse	68
8.2.1. Überprüfung der Hauptkomponentenmethode	68
8.2.2. Informationstheoretische Reduktion der Dimensionen	70
8.3. Programmaufbau	74
9. Schlußbemerkung	79
Literaturverzeichnis	81

0. Übersicht

Gegenstand dieser Arbeit ist die Konstruktion und Erprobung eines Verfahrens zur Selektion von Krankheitssymptomen aufgrund der enthaltenen Trenninformation.

Das informationstheoretische Verfahren wird im Zusammenhang mit den statistischen Methoden zur Datenkomprimierung und -klassifikation gesehen, wobei die direkte Anwendung in der Medizin im Vordergrund bleibt.

Mit Hilfe eines von Kullback entwickelten informationstheoretischen Formalismus wird die Hauptkomponentenmethode im Rahmen der Informationstheorie diskutiert und ihre Effizienz praktisch überprüft.

Die Anwendung mathematischer Verfahren in der Medizin befindet sich in einem Anfangsstadium, wobei bisher keinerlei konkrete, universal anwendbare Methoden vorliegen. Die Frage, die sich die Forschung stellt ist, ob der komplexe Diagnoseprozess des Arztes mit Hilfe von automatischen Verfahren nachvollziehbar ist, und wenn ja, welche der bekannten mathematischen Algorithmen am besten dafür geeignet sind. Um diese Frage zu beantworten, ist sowohl eine gründliche Analyse des üblichen Diagnoseprozesses erforderlich, als auch die Entwicklung neuer, dem Zweck angepaßter, mathematischer Verfahren.

Aufgrund dieser Tatsachen wird zunächst ein Forschungskonzept entwickelt, aus welchem sich die einzelnen Schritte dieser Untersuchung herleiten und begründen lassen.

Die Arbeit läßt sich daher in drei Teile untergliedern:

(1) Einführender Teil

Die Kapitel 1. bis 3. enthalten die Beschreibung des Diagnoseprozesses, die chronologische Entwicklung, sowie die Problemstellung der computerunterstützten Diagnose.

Im Kapitel 4. werden die wichtigsten Klassifikationsverfahren und die neuesten Standpunkte der Klassifikationstechnik in der Medizin dargestellt. Konkrete geometrische Interpretationen ergänzen das Spektrum der abstrakten Klassifikationsverfahren.

Der Zusammenhang zwischen der Anzahl der Symptome, dem Umfang der bekannten Patientendaten und der Fehlklassifikationsrate ist im Kapitel 5. enthalten. Relevant für das Ziel dieser Arbeit ist hier insbesondere der Einfluß der Symptomanzahl auf die Fehlklassifikationsrate. Eine Diskussion der bekannten Verfahren zur Reduktion der Symptome, liegt im Kapitel 6. vor.

(2) Theoretischer Teil

Eine informationstheoretische Betrachtungsweise (Kullback) wird hier (Kapitel 7.) zur Erklärung der Hauptkomponentenmethode verwendet. Mit Hilfe des informationstheoretischen Formalismus versucht man dabei die Einschränkungen dieser Methode auszuprägen.

Eine Erweiterung des Formalismus und die Konstruktion einer Methode zur Reduktion der Symptome aufgrund der enthaltenen Trenninformation (Variablendivergenz) wird vorgenommen.

(3) Empirischer Teil

Die Planung und der Ablauf der Simulationsuntersuchungen werden im Kapitel 8. beschrieben. Mit Hilfe eines selbstgebaute Datenzufallsgenerators werden zunächst pseudozufällige Vektoren (Patientendaten) erzeugt. Es

folgt eine Verarbeitung dieser künstlich erzeugten Patientendaten mit Hilfe von statistischen Programmen. Zu diesem Zweck wird ein Programm zur informationstheoretischen Reduktion der Dimensionen entwickelt und in das, aus Darmstadt stammende statistische Programm DISK (Programm für multiple Diskriminanzanalyse und Zuordnungsverfahren) implementiert.

Die Monte-Carlo- Experimente werden zur Überprüfung der theoretischen Aussagen vom Kapitel 7. verwendet.

Das Hauptgewicht dieser Arbeit liegt auf der informationstheoretischen Betrachtungsweise bei der Selektion der entscheidendsten Symptome in der computerunterstützten Diagnose. Dazu zählt man die Konstruktion und Erprobung eines Algorithmus zur Reduktion der Symptome.

Ein wesentliches Ziel dieser Arbeit sehe ich darin, eine minimale Fehlklassifikationsrate durch Datenkomprimierung zu erreichen.

1. Einleitung

Die ärztliche Behandlung eines Patienten ist ein zeitabhängiger, dynamischer Prozeß, der aus der Feststellung der Diagnose, aus der Beobachtung des Krankheitsverlaufes und der Wirkung unterschiedlicher Therapien besteht. Um einen solchen komplexen Prozeß mit Hilfe des elektronischen Rechners unterstützen zu können, muß man ihn definieren, unterteilen und mit Hilfe eines mathematischen Modells beschreiben.

Im Rahmen einer mathematischen Analyse ist zuerst ein statisches Bild des Diagnoseprozesses erforderlich. In diesem Fall ist die ärztliche Diagnosefindung eine Voraussage der möglichen Diagnosen anhand gespeicherter Daten (über Krankheiten einer bestimmten Population), wenn ein akuter Fall mit einer gewissen Symptomkombination vorliegt (Abb. 1). Der Begriff "Voraussage" bezieht sich auf das Ziel der Diagnosefindung: die Früherkennung einer Krankheit zu ermöglichen, um irreversible Zustände des betroffenen Organismus zu vermeiden.

Als grundlegende Elemente der folgenden Analyse, müssen die Begriffe: Diagnose, Krankheit, Symptom und Diagnoseprozeß definiert werden.

Die Krankheit - Ursache der Diagnosefindung -, kann als ein abnormaler, ggf. pathologischer Zustand des menschlichen Organismus angesehen werden. Pathologische Fälle, die eine ähnliche Konstellation der Symptome, einen ähnlichen Verlauf des abnormalen Zustandes und eine ähnliche Reaktion

bei gleichen Therapien aufweisen, werden einer statistischen Gruppe zugeordnet. (Die Begriffe: Gruppe, Krankheit, Population, werden im weiteren als synonym verwendet.) Solche Gruppen werden mathematisch durch die Zusammenstellung der bereits diagnostizierten Patienten (erhärteten Fälle) bestimmt, wodurch eine Stichprobe des untersuchten statistischen Phänomens gebildet wird.

Die Krankheiten können mittels der spezifischen Symptomkonstellation erkannt werden. Die Symptome sind Anzeichen des krankhaften Zustandes des Organismus [18].

Um eine sinnvolle Analyse des Diagnoseprozesses durchführen zu können, müssen alle Merkmale berücksichtigt werden, die zur Identifikation einer Krankheit führen. Dazu gehören allgemeine Personaldaten (z.B. Alter, Geschlecht), Labor-tests (z.B. das O_2 -Äquivalent im Blut), Krankheitssymptome (z.B. Fieber, Schmerzen) und Befunde. Für jede Krankheit werden alle typischen Merkmale in einer Liste zusammengefaßt, die auch eine Variablenliste der statistischen Aufgabe darstellt. Werden bei einem neuen Patienten bestimmte Merkmale festgestellt, weist man den entsprechenden Variablen bestimmte Werte zu.

(Im folgenden werden Merkmale und Symptome als synonym verwendet)

Patienten die an eine und derselben Krankheit leiden, haben meistens eine ähnliche Symptomkombination (Krankheitsbild). Man kann behaupten, daß jede Krankheit eine gewisse Liste von typischen Symptomen bestimmt, und umgekehrt, kann eine typische Konstellation von Symptomen als Krankheit bezeichnet werden, auch wenn man ihre gemeinsame Ursache nicht kennt [18].

In dieser Arbeit werden die Symptome einer Krankheit als zufällige Variablen der entsprechenden statistischen Gruppe betrachtet. Diese Annahme ist insofern berechtigt, da das Auftreten jedes Symptoms mit Unsicherheit verbunden

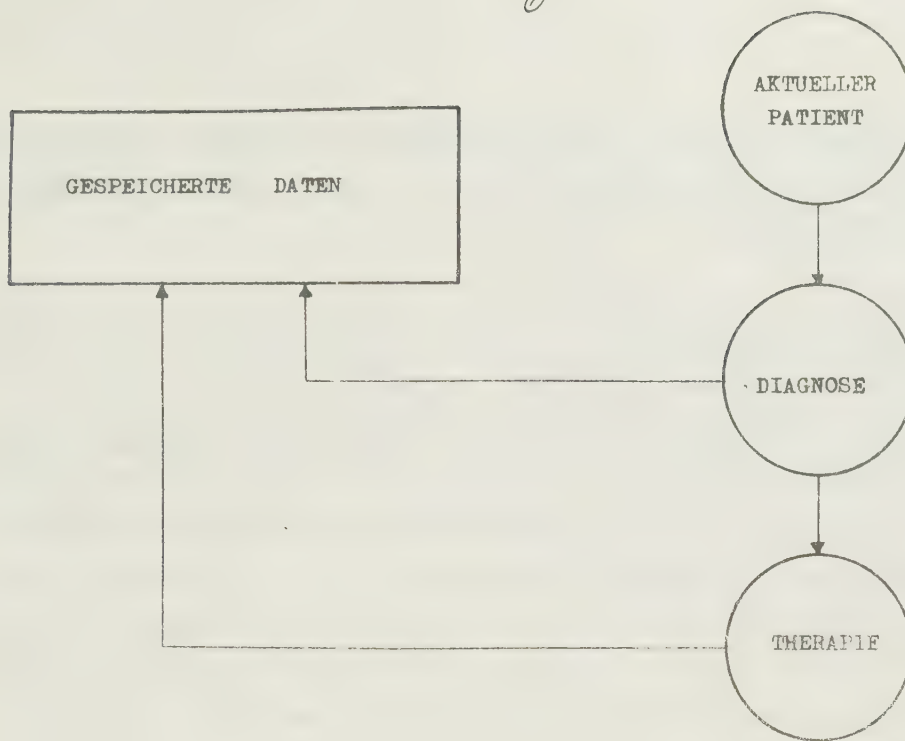


Abb.1 - Blockbild der statischen Diagnostik

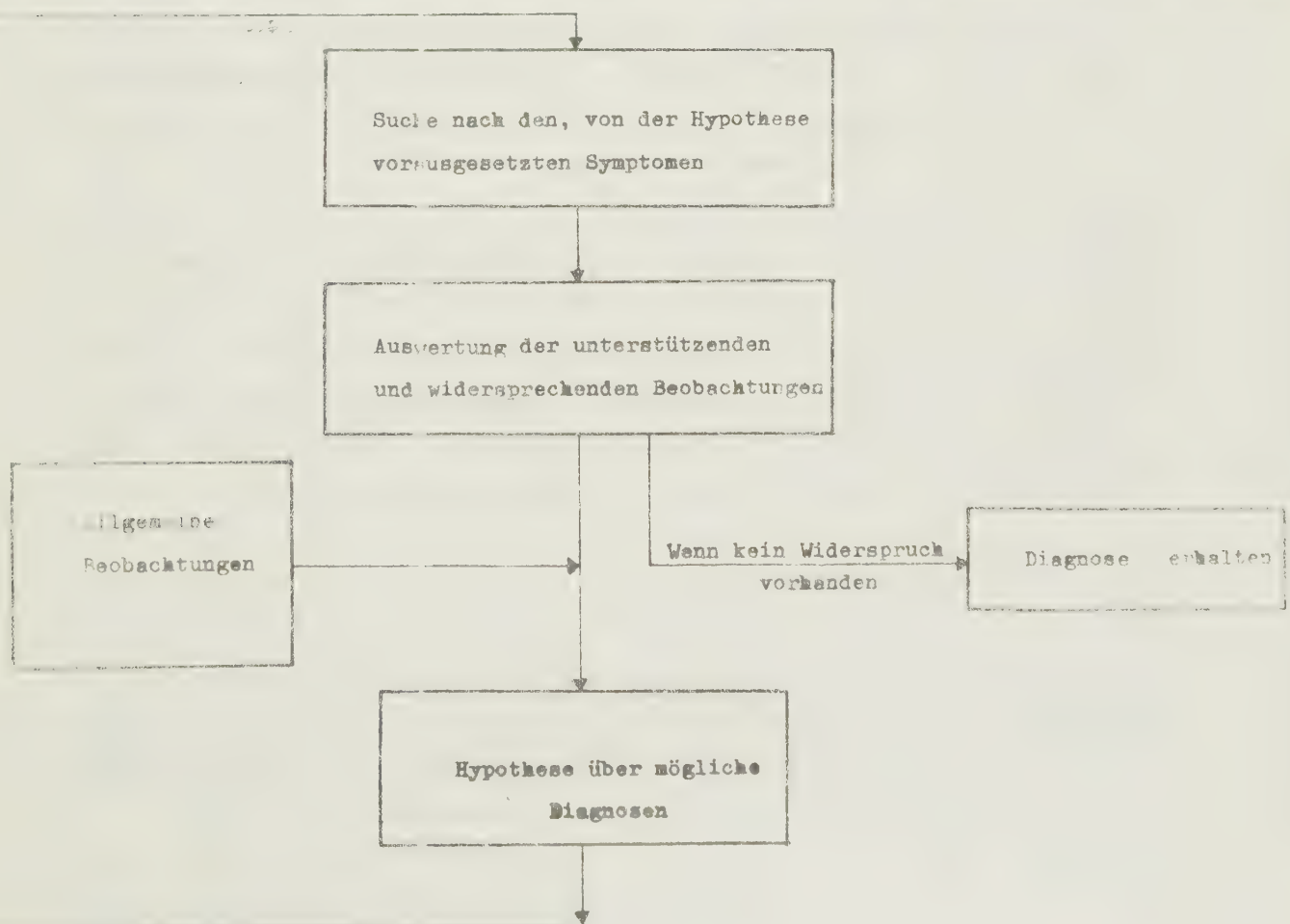


Abb.2 - Flußdiagramm der sequentiellen Diagnostik

ist. Die gesamte Anzahl der Variablen bestimmt die Dimension des Problems. Hierbei werden alle Symptome der untersuchten Krankheiten berücksichtigt.

Die Variablen können diskret oder stetig sein, je nach dem, ob die entsprechenden Symptome diskret, bzw analog ausgewertet werden können. Einige Labortests und die Feststellung von Blutdruck, Fieber, Pulsfrequenz etc, ergeben Meßdaten, die am besten durch stetige Funktionen dargestellt werden können. Für andere Symptome wie z.B. Husten, verschiedene Schmerzen, radiologische Feststellungen etc, gibt es nur die beiden Werte "vorhanden" oder "nicht vorhanden", d.h., sie können nur binär ausgewertet werden.

Das Endziel des Diagnoseprozesses ist die Diagnose. Das bedeutet, daß im Falle eines neuen Patienten, die vorliegende Krankheit mit absoluter Sicherheit erkannt werden muß. Häufig kann das diagnostische Ergebnis nicht bei den ersten Versuchen festgelegt werden. Zunächst geht man von einer Hypothese aus. Das setzt eine Untersuchungsstrategie voraus, die zur Entdeckung weiterer Symptome führt. Aufgrund der nicht festgestellten Symptome kann die Hypothese bestätigt oder zurückgewiesen werden. Falls die erste Hypothese zurückgewiesen wurde, folgt eine zweite Hypothese. Danach sucht man wieder nach spezifischen Symptomen, usw. Dieses Vorgehen definiert einen zeitabhängigen Diagnoseprozeß. Krischer [32] beschreibt den Diagnoseprozeß in einem Flußdiagramm, das in der Abb. 2 dargestellt ist.

Es gibt sehr wenige Symptome die eine hypothetische Diagnose mit absoluter Sicherheit bestätigen oder ablehnen können. Oft werden keine neuen Merkmale mehr gefunden, die die Hypothese verstärken. In diesem Fall sind Versuchstherapien zu empfehlen. Ihre Wirkung auf den Krankheitsverlauf wird für oder gegen die Hypothese entscheiden.

Ein Ziel der automatischen Diagnosefindung ist, den Diag-

noseprozeß schnell konvergent zu machen und den interaktiven Eingriff des Arztes in den Prozeß zu ermöglichen.

Die ersten Versuche, ein mathematisches Diagnosemodell aufzubauen, wurden mit dem Ziel durchgeführt, der ärztlichen Logik bei der Diagnosefindung nachzugehen.

Das logische Modell [37] [50] setzt sich aus dem Satz der berücksichtigten Diagnosen (Krankheiten) $D = [D_1, \dots, D_q]$

und dem Satz der dafür entscheidenden Symptome $S = [S_1, S_2, \dots, S_m]$ durch Boole'schen Funktionen zusammen.

Dabei sind die Symptome S_i und die Diagnosen D_i durch logische Operatoren verknüpft. D_i und S_i werden als logische Variablen betrachtet, denen ein Wert "wahr" oder "falsch" zugewiesen wird.

Die Vielzahl der von der möglichen Symptomkombinationen abgeleiteten logischen Funktionen macht eine komprimierte Datenerfassung unmöglich. Dadurch ist die Anwendung des logischen Modells mit Schwierigkeiten verbunden.

Weil jeder neue Patient eine zufällige Symptomkombination aufweist, und die ärztliche Diagnose selbst meistens mit Unsicherheit verbunden ist, könnte ein statistisches Modell der Realität besser angepaßt werden. Die Wahrscheinlichkeitstheoretischen Modelle [37] haben auch den Vorteil, die Unsicherheit einer Diagnose mit Hilfe von Wahrscheinlichkeitsmaßen darstellen zu können. Falls absolute apriori Wahrscheinlichkeitsmaße $P(D_i)$ für das Auftreten der Krankheiten bei der berücksichtigten Population (wenn z.B. Epidemien, Betriebskrankheiten, etc. vorkommen) bekannt sind, werden meistens im automatischen Diagnoseprozeß Bayes'sche Modelle eingesetzt. Man erhält hier ein aposteriori Wahrscheinlichkeitsmaß $P(D_i / S_1, S_2, \dots, S_m)$ für die einzelnen Diagnosen D_i unter Voraussetzung [46] eines gegebenen Symptomkomplexes $S = [S_1, S_2, \dots, S_m]$.

Oft erfolgt die Herleitung der Entscheidungsregel direkt vom Likelihood- Quotienten- Kriterium (die einfachste Form der Bayes- Strategie) über Diskriminanzfunktionen.

Für den Fall, daß keine apriori Wahrscheinlichkeiten vorliegen, ist die Bayes- Strategie nicht anwendbar. Hier können spieltheoretische Modelle in der medizinischen Diagnosefindung eingesetzt werden. Im Sinne der Entscheidungstheorie werden dabei Risikofaktoren für die Fehldiagnose definiert, die mit der Patiententherapie verbunden sein können. Man versucht dann, unter allen zulässigen Entscheidungsverfahren jenes herauszufinden, das den größtmöglichen erwarteten Verlust minimiert. Dieses Vorgehen ist auch als Minimax- Lösung bekannt.

2. Chronologische Entwicklung der mathematischen Methoden zur Diagnoseunterstützung

Um einige der grundlegenden Fragen in den bisherigen Methoden zur maschinellen Entscheidungshilfe hervorzuheben, wird im folgenden ihre Entwicklung skizziert.

Mathematische Modelle zur Diagnoseunterstützung wurden zum ersten Mal von Ledley und Lusted [37] im Jahre 1959 diskutiert. In ihrer Arbeit "Reasoning Foundations of Medical Diagnosis" beschreiben sie die drei möglichen mathematischen Modelle: das logische, das wahrscheinlichkeitstheoretische und das spieltheoretische Modell als Grundlage für den Computereinsatz in dem ärztlichen Diagnoseprozeß.

1963 teilten Rinaldo u.a. [44] die Symptome in zwei Sätze auf: in einem Satz mit Symptomen die einen wesentlichen Beitrag zur Trennung der Krankheiten bringen und in einem Satz mit den restlichen Symptomen.

Es folgen eine Reihe von Arbeiten, die aufgrund des Bayes'schen Theorems und unter der Voraussetzung voneinander unabhängiger Symptome, eine Differenzialdiagnostik für konkrete Fälle erprobten. So kann z.B. die Diagnose angeborener Herzkrankheiten, durchgeführt von Warner u.a. (1961) [51], erwähnt werden. Die Bayes'sche Formel wurde auch von Scheinok und Rinaldo (1967) [47] zur Diagnoseunterstützung im Falle überabdominaler Schmerzen verwendet.

Ein Jahr später versuchten dieselben Autoren [48] ein anderes mathematisches Verfahren - die Diskriminanzanalyse-

- 1 -

für dieselbe Diagnostik anzuwenden. 1971 entstand ein Vergleich [49] zwischen den beiden Methoden. In den Experimenten wurden sechs Krankheiten untersucht : Hiatus-Hernie, Zwölffingerdarmgeschwür, gastrischer Geschwür (Magengeschwür), Krebs, Gallensteine und Funktionskrankheit. Eine Liste von 11 Symptomen ist dabei berücksichtigt worden. Für 224 von 300 Patienten ergaben die zwei mathematischen Modelle (Bayes'sche Verfahren und Diskriminanzanalyse) dasselbe Ergebnis; es wurde jedoch nur in 144 Fällen (48%) die richtige Diagnose gestellt.

Andere Versuche wurden mit derselben Zielsetzung von Gorry 1968 [26], Reale u.a. (1968) [43], Bouckaert (1972) [10], Bouckaert u. Beckers (1973) [11] durchgeführt. Die geringe Prozentzahl der korrekten Zuordnungen - eine typische Eigenschaft dieser Ergebnisse - ist meistens wegen der Mißachtung der Abhängigkeiten zwischen den Symptomen zu erklären. Die Bayes'sche Formel wurde unter der Voraussetzung angewandt, daß unterschiedliche Symptome voneinander statistisch unabhängig sind.

Diese Versuche stießen auf Schwierigkeiten wegen der Auswertung von wenigen erhärteten Fällen. Die erforderlichen Stichproben verfügen meistens über eine geringe Anzahl der bereits diagnostizierten Fälle und genügen deshalb nicht einer relativ genauen Abschätzung der Verteilungsparameter. Die Auswirkung der Parameterschätzung von nichtausreichenden Stichproben ist zum ersten Mal unter medizinischem Aspekt von Bailey (1965) [5] berücksichtigt worden. Er hat eine neue Schätzungsmethode für die multinomiale Verteilungsfunktion vorgeschlagen, für den Fall, daß kleine Stichprobenmengen vorliegen.

Gorry und Barnet (1968) [27] betrachteten zum ersten Mal die Diagnostik als einen dynamischen Prozeß. Das führte zu sequenziellen Diagnosestrategien. 1970 wurde eine sequen-

tielle Zeichenerkennungstechnik für die medizinische Diagnose von Lasker [36] vorgeschlagen. Damit wurde eine Untersuchungsstrategie entwickelt, mit dem Ziel, in möglichst kurzem Zeitraum die zur Trennung der Krankheiten am besten geeigneten Symptome zu finden.

Diese neuen Gesichtspunkte haben das Konzept der Diagnosestrategie erheblich geändert. Das kann man deutlich im Krischers (1971) [32] Flußdiagramm des heuristischen Diagnoseprozesses sehen (Abb.2)

Um die notwendigen Anforderungen des Bayes'schen Modells bezüglich der Unabhängigkeit zwischen Symptomen zu vermeiden, sind auch deterministische Algorithmen zur Diagnoseunterstützung entwickelt und erprobt worden. In diesem Zusammenhang können die Arbeiten von Cumberbatch und Heaps (1973), (1974) [15] [16] erwähnt werden. Es handelt sich im wesentlichen um polinomial Modelle, wobei die Terme größeren Grades als eins, der Abhängigkeit zwischen den Symptomen entsprechen. Da meistens die Terme dritten Grades bei der Berechnung vernachlässigt werden müssen, stößt man hier auf gleiche Schwierigkeiten.

Innerhalb der Entwicklung automatischer Diagnoseverfahren muß die Kollektion von biomedizinischen Programmen, herausgegeben von Dixon (1970) [17], als ein wertvoller Forschungsbeitrag betrachtet werden.

Charakteristisch für die bisher geschilderten Anwendungsversuche der maschinellen Diagnoseunterstützung ist die isolierte Bearbeitung diagnostischer Teilbereiche. Die automatischen Verfahren könnten nur für Gruppen von ähnlichen Krankheiten (die meistens mit einem bestimmten Körperteil oder Organ verbunden sind) eingesetzt werden. Hinzu kommt, daß die Anwendung dieser Methoden eine enge Zusammenarbeit unterschiedlicher Spezialisten erfordert um die medizinischen Erfahrungen zusammenzutragen, die Programme zu

schreiben und den Einsatz des Diagnosesystems zu ermöglichen. Das bedeutet einen Arbeitsaufwand von vielen Mann-Jahren und eine Steigerung der Entwicklungskosten.

Dem Anwender wurden Diagnosesysteme angeboten, die ihm die Symptombewertungen abnahmen und Diagnoseaussagen lieferten, deren Beschränkungen aber, er nicht überschauen konnte. Unter solchen Bedingungen ist es verständlich, das die Ärzte ihnen fremde, fast selbständige Computer-Diagnosesysteme ablehnten. Die Probleme werden weit größer, wenn unterschiedliche diagnostische Bereiche von einem Zentralen Diagnosesystem parallel unterstützt werden sollen.

Ein wichtiger Schritt, diese Schwierigkeiten zu überwinden, wurde durch die Entwicklung eines globalen Informationssystems zur Diagnoseunterstützung in der Bundesrepublik Deutschland unternommen. Hölzel [29] hat 1973 in der Abteilung für medizinische Statistik, Dokumentation und Datenverarbeitung an der Universität Ulm den Anfang gemacht. Das System besteht aus einer Programmsammlung, die eine einfache Speicherung von Daten unabhängig von dem diagnostischen Bereich ermöglicht und ihre Auswertung durch Diagrammen, Diagnosealgorithmen, Histogrammen usw. erleichtert. Aufgrund des Dialoges Arzt-Maschine, kann das System eine schrittweise durchgeführte Diagnosestrategie unterstützen. Die theoretische Grundidee der maschinellen Diagnoseunterstützung wurde dadurch der Praxis angenähert. Die Algorithmen selbst sind aber unverändert geblieben.

Ein Vergleich zwischen der prognostischen und der mit Hilfe der geschätzten Parameter ermittelten Diagnose entstand 1973 in der Arbeit von Aitchison und Kay [1]. Hier versuchte man erneut die Fluktuation der geschätzten Parameter in Abhängigkeit von der Größe der Stichprobe zu vermeiden. Dafür wurden prognostische Verteilungsfunktionen verwendet und ihr Einsatz in der klinischen Diagnosefindung studiert.

Einen wichtigen Beitrag zur Untersuchung statistischer Verfahren zur Diagnoseunterstützung hat der GSF-Bericht MD 95 [12] gebracht. Christl und Stock, die Autoren dieses Berichtes, haben drei spezielle, auf der statistischen Methode basierenden Verfahren, in die Untersuchung einbezogen. Dazu gehört einmal das Bayes'sche Verfahren, das in den bisher verwendeten Algorithmen eine führende Rolle gespielt hat. Zum anderen sind die unter dem Namen "Diskriminanzanalyse" bekannten Verfahren bei der Untersuchung berücksichtigt worden (der lineare und der quadratische Ansatz). Merkmale der sowohl diskreten, als auch stetigen Verteilungstypen wurden über einen modular aufgebauten Zufallszahlgenerator erzeugt. Dadurch konnten Leistungsfähigkeit und Güte der obenbeschriebenen Methoden untersucht werden. Der Einfluß von systematischen und zufälligen Fehlern bei der Symptomerhebung auf die Leistungsfähigkeit der Verfahren stand im Mittelpunkt dieser Untersuchungen.

Neue und in der Medizin nützliche Methoden zur Abbildung eines m -dimensionalen Entscheidungsraumes auf eine Ebene sind durch die Arbeiten von Andrews [3] und Sammon [45] bekannt geworden. Ein neuer medizinischer Diagnoseautomat wurde mit Hilfe einer räumlichen Darstellung der statischen Diagnostik von Balanescu [6] beschrieben.

Weitere Arbeiten, wie z.B. die von Geisser (1964) [22], Lachenbruch (1965) [35], Hills (1966) [28], Cox (1970) [14], Gilbert (1968) [23], Martin u. Bradley (1972) [38], Blomer u. Drechsler (1976) [9] und andere, haben einen wichtigen Beitrag zur Entwicklung automatischer Diagnosealgorithmen geleistet.

3. Auflistung der Fragestellungen

Eine Aufgabe der automatischen Diagnoseunterstützung ist unter anderem, die Beantwortung der Frage, inwieweit die Diagnosefindung des Arztes mit formalen Methoden nachvollzogen werden kann. Trotz einer mehr als 15-jährigen theoretischen Erfahrung auf diesem Gebiet, gibt es noch viele wichtigen Fragen zu beantworten. Die sporadischen praktischen Versuche haben einige medizinische Anforderungen verdeutlicht. Wenige davon sind von den bisher bekannten mathematischen Algorithmen erfüllt. Die ungelösten Probleme können unter drei Aspekten berücksichtigt werden, die sich auch aus der geschichtlichen Entwicklung der Diagnoseunterstützung ergeben:

- (1) Die erforderliche medizinische Erfahrung (erhärteten Fälle) zur Definition der Krankheiten muß einfach und ohne übertriebenem Personalaufwand erfolgen. Obwohl die Informationssysteme zur Diagnoseunterstützung viele Aufgaben in dieser Hinsicht erleichtert haben, bleiben bestimmte Fragen immer noch offen.

Erstens, darf der Aufbau der Datenbank keine zu großen Stichprobenmengen erfordern. Oft ist es schwierig die erforderliche Anzahl der bereits diagnostizierten Patienten mit gewissen Krankheiten (z.B. die Lungenkrankheit 'Aktinomykose') in einem Krankenhaus zu finden. In solchen Fällen hat man nur eine kleine Stichprobe solcher Patienten zur Verfügung, die gleichzeitig die Größe der anderen Stichproben bestimmt (gleich große Stichproben für jede Krankheit vereinfacht den mathematischen Formalismus). Um die Parameter der statistischen Verteilungsfunktionen aus kleinen Stichprobenmengen schätzen zu können, müssen

bestimmte Methoden ausgesucht und standardisiert werden (vgl. [5]). Das vereinfacht die Datenerfassung und reduziert dadurch die Arbeit, die Datenbank zu erstellen.

Zweitens, ist es wichtig, daß bei der Zusammenstellung der erhärteten Fälle kein hochqualifiziertes Personal notwendig ist. Obwohl dialogfähige Informationssysteme das Sammeln von Merkmalen und Aus- und Bewertung im wesentlichen erleichtert haben, ergeben sich immer noch ungelöste Aufgaben. So müssen z.B. die mit der Erkrankung verbundenen apriori Informationen (über Epidemien, ungesunde Arbeits- und Lebensbedingungen, etc.) extra notiert werden. Die davon betroffenen Patientendaten werden in die Berechnung der apriori Wahrscheinlichkeiten einbezogen. Die Wirkung unterschiedlicher Therapien auf bestimmte Krankheitsfälle muß auch im voraus bekannt sein, wenn in der sequentiellen Diagnostik Versuchstherapien verwendet werden sollen.

Eine vollständige Symptomliste der untersuchten Krankheiten ist erforderlich. Welche Symptome oder Syndrome bei der Untersuchung bestimmter Krankheiten zu berücksichtigen sind, muß von Anfang an der Arzt entscheiden. Dazu stellt er auch die mögliche Auswertung (diskret oder stetig) der Symptome fest. Manche Symptome, die sich meist stetig verhalten, z.B. die Blutsenkungsgeschwindigkeit, können als diskrete Symptome interpretiert werden. (BSG normal, BSG verzögert, BSG beschleunigt). Die formalen Methoden für die Datenerfassung sind immer noch nicht geeignet, solche Auswertungen zu übernehmen. Die Zusammenarbeit der Ärzte mit den Systemingenieuren bleibt hier die einzige Lösung.

Diejenigen Merkmale, die am meisten zur Unterscheidung der Krankheiten beitragen können, aus der gesamten Syptomliste herauszufinden, bleibt hier ein weiteres Problem. Oft ist es notwendig, die Dimension

des Entscheidungsproblems zu reduzieren, um den Rechenaufwand in einem erforderlichen Verhältnis zu den Rechenanlagen zu halten. Welches von den bekannten Verfahren ist dafür am besten geeignet? Kann man mit Hilfe einer Reduktion der nichtrelevanten Symptome die verschiedenen Krankheiten besser voneinander unterscheiden?

Folglich, kann man die Behauptung aufstellen, daß alle Diagnosealgorithmen eine Komprimierung und Abstraktion der medizinischen Erfahrung erfordern.

- (2) Bisher sind mehrere Algorithmen zur Diagnoseunterstützung bekannt. Die meisten folgern aus den zwei möglichen Lösungen eines Entscheidungsproblems: die Bayes'sche Lösung und die Minimax-Lösung (s. §4.3). Varianten für stetige und diskrete Variablen (Merkmale) sind aus diesen zwei Hauptverfahren abgeleitet worden. Erwähnenswert sind hier auch die graphischen Diskriminanzmethoden [3] [45], die eine wertvolle Hilfe in der klinischen Diagnostik leisten können. Es ist weniger bekannt, welche davon sich am besten für ein einheitliches Diagnosesystem eignet. Die bisherigen Versuche, einen Vergleich bekannter Methoden durchzuführen, lassen noch keine abschließende Beurteilung zu. Es sind keine, sich daraus ergebenden Ansätze für die automatische Diagnoseunterstützung vorhanden.

Ein weiteres Problem ist das Verhalten solcher Algorithmen im Falle voneinander abhängiger Symptome. Die Probleme werden weit größer, wenn gleichzeitig stetige und diskrete Symptome bewertet werden. Welches von den bekannten Verfahren ist in diesem Fall für eine rasche und effiziente Diagnosefindung geeignet?

Viele dieser Fragen können nur durch Experimente beantwortet werden. Aus dem Kapitel 2. könnte man schließen, daß einige Methoden in konkreten pathologischen Fällen schon geprüft worden sind. [47], [11]. Solche Versuche berücksichtigen aber meistens nur isolierte diagnostische Teilbereiche. Sie sind mit einem großen Zeit- und Arbeitsaufwand verbunden und reichen deshalb für einen vollständigen Methodenvergleich nicht aus. Die pseudo-zufälligen Experimente (Monte-Carlo-Experimente) auf digitalen Rechnern haben sich von diesem Gesichtspunkt als günstiger erwiesen [12].

- (3) Die Konvergenz des Diagnoseprozesses ist ein weiterer Aspekt der Diagnoseunterstützung. Ein effizientes Diagnosesystem muß eine schnellere Konvergenz bieten können, als es der Arzt bei der Diagnosefindung bisher erreicht hat. Was bestimmt unter diesen Umständen die Prozessschritte? Welche sind die grundlegenden Kriterien, die zum Aufbau einer optimalen Untersuchungsstrategie führen?

Eine Aufteilung des Diagnoseprozesses in diagnostische Etappen (wie z.B. 1.Etappe: Anamnese, 2.Etappe: ärztliche Untersuchung, 3.Etappe: radiologische Untersuchung, 4.Etappe: Labortest usw.) bringt eine gewisse Konvergenz, die aber nicht immer optimal ist. Wählt man für den nächsten Untersuchungsschritt die Symptome gezielt aufgrund von wahrscheinlichkeitstheoretischen Kriterien, so wird die Konvergenz verbessert. Zwei Beispiele aus der Literatur sind dafür repräsentativ. Nach einer Zeichnungserkennungstechnik, 1970 von G.E. Lasker [36] vorgeschlagen, sind diejenigen Symptome auszuwählen, die eine maximale Häufigkeitsrate für die hypothetische Krankheit aufweisen. W.J.Dixon (1970) [17] dagegen, teilt den Symptomensatz in zwei Untersätze auf: zu einem Unter-

satz gehören einmal die Symptome die in den nächsten Entscheidungsschritt einbezogen werden, zum anderen gehören die übrigen Symptome. Es folgt eine schrittweise durchgeführte Kovarianzanalyse, die das Kriterium für die Übernahme eines Symptoms von einem Untersatz zum anderen liefert. Es liegt keine praktische Prüfung dieser Methoden vor, was eine Entscheidung für die eine oder andere Methode erschwert.

Aber selbst wenn am Ende einer sequentiellen Untersuchungsstrategie eine Diagnose erstellt wird, ergeben sich neue Fragestellungen. Welche Therapie ist für den Patienten am besten geeignet? Welche Medikamente schaden seinem Organismus nicht? Welche Therapie eignet sich am besten, schnellstens die festgestellte Diagnose zu bestätigen, ohne, daß im Falle einer Mißdiagnose dem Patienten geschadet wird? Hier sind neue Patientendaten erforderlich. Es ist notwendig für die jeweilige untersuchte Krankheit, die Wirkung jeder Therapie in Tabellen statistisch zu erfassen. Risiken der Therapien für jede mögliche Krankheit, sind dabei zu definieren.

Aus den obenbeschriebenen Problemen ist zu schließen, daß das aktuelle Stadium der automatischen Diagnoseunterstützung immer noch keine vielseitige und sichere Anwendung bietet. Wie D. Hölzel ([29] S.16) dazu meint, "Wird all diesen Problemen die Tatsache gegenübergestellt, daß in der überwiegenden Anzahl der zu behandelnden Fälle eine maschinelle Entscheidungshilfe nicht notwendig ist, wegen des erforderlichen Zeitaufwandes nicht realistisch ist, so wirkt die Behauptung, daß mit den bisherigen Studien der Nachweis einer praktischen Nutzung erbracht wurde, provozierend und unglaubwürdig zugleich. Dieser von allzueinseitiger technologischer Mentalität getragene hohe Anspruch ist aufzugeben. Er ist eine Fiktion. Primär geht es darum,

die vielseitigen und sinnvollen Anwendungsmöglichkeiten von Computern aufzuzeigen, mit denen der Medizin als Erfahrungswissenschaft Rechnung getragen und wesentliche Unterstützung angeboten werden kann. Dabei bleibt das Ziel bestehen, nämlich die Entwicklung diagnostischer Modelle, mit denen der Diagnoseprozess des Arztes nachvollzogen und somit zur Aufbereitung von Entscheidungsgrundlagen benutzt werden kann."

In der vorliegenden Arbeit wird folgende Phase des automatischen Diagnoseprozesses untersucht: die Minimierung der Fehldiagnose durch Selektion der zur Unterscheidung der Krankheiten relevanten Symptome. Innerhalb dieser Phase beschränken sich die theoretischen und experimentellen Untersuchungen auf stetige Symptome.

4. Klassifikationsverfahren zur Diagnoseunterstützung

4.1. Die Problemstellung

Geht man davon aus, daß ein bestimmter diagnostischer Teilbereich ausgewählt wurde, so können die damit verbundenen Krankheiten $\{ D^k, k = 1, 2, \dots, q \}$ ermittelt werden. Weiterhin werden diese Krankheiten als statistische Gruppen (Populationen) betrachtet. Um die Gruppen aus der Gesamtheit der Patientendaten zu bestimmen, werden zunächst Stichproben mit bereits korrekt diagnostizierten Patienten gebildet. Eine Stichprobe stellt die gemeinsamen Eigenschaften einer Population zusammen und bildet dadurch eine Klasse oder ein "Klassifikationsgebiet". Alle diese Klassen sind mittels einer Merkmalliste $s_1, s_2, \dots, s_m$ zu beschreiben.

Eine statistische Entscheidungsaufgabe besteht darin, aufgrund der festgestellten Symptome eines Patienten zu entscheiden, welche Krankheit (Klasse) vorliegt. Die statistischen Verfahren zur Bestimmung einer medizinischen Diagnose unterscheiden sich von den logischen Verfahren grundsätzlich dadurch, daß das Merkmalmuster eines Patienten

$$(4.1) \quad s^* = \{ s_a, s_b, \dots, s_l \} ; a, b, \dots, l \in \{ 1, 2, \dots, m \}$$

als Zufallsgröße aufgefaßt wird. Ein Patient mit einem konkreten Merkmalmuster ist dann eine Realisation dieser Zufallsgröße, die einer bestimmten Verteilung gehorcht.

Unter diesen Voraussetzungen ist es möglich, einen Stichprobenraum

$$(4.2) \quad S = \{ s_i, i = 1, 2, \dots, m \}$$

im Sinne der Wahrscheinlichkeitstheorie zu definieren, wo jede Realisation als ein Punkt dargestellt werden kann.

Um weiter das Problem in die mathematische Sprache zu überführen, werden die Zufallsvariablen

$$(4.3) \quad X_i: S \rightarrow E_1 \quad ; i = 1, 2, \dots, m$$

definiert, wobei E_1 eine reelle Achse ist. Eine Zufallsvariable entspricht in diesem Fall einem Merkmal und ist eine Abbildung des Stichprobenraumes auf die reelle Achse. Es wird dabei vorausgesetzt, daß die Zufallsvariablen einer multivariaten Verteilung gehorchen, die durch eine Dichtefunktion $f(X_1, X_2, \dots, X_m)$ beschrieben wird. In dem Fall, daß eine bestimmte Verteilungsfunktion zugelassen ist, wird sie als für jede Krankheit unterschiedlich angenommen und repräsentiert diese Krankheit. Das wird durch einen unteren Index dargestellt (z.B. $f_k(X_1, X_2, \dots, X_m)$ wenn die Verteilungsfunktion der Krankheit D^k entspricht).

Die oben definierten Zufallsvariablen werden im folgenden mit Hilfe des m -dimensionalen Zufallsvektors $\mathbf{X}^T = [X_1, X_2, \dots, X_m]$ zusammengefaßt. Hierbei und im Laufe dieser Arbeit wird mit dem oberen Index T die Transponierte einer Matrix (eines Vektors) bezeichnet.

4.2. Bayes - Lösung

Geht man davon aus, daß ein Patient durch m Merkmale charakterisiert wird, so kann das Merkmalmuster (4.1) als Realisation $\mathbf{X} = [x_1, x_2, \dots, x_m]$ des m -dimensionalen Zufallsvektors

$$(4.4) \quad \mathbf{X}^T = [X_1, X_2, \dots, X_m]$$

betrachtet werden. Gesucht ist die Klassifikation einer solchen Beobachtungsgröße $\mathbf{X}$ (eines Individuums, für das gilt $\mathbf{X} = \mathbf{x}$) in eine von q verschiedenen vorgegebenen Populationen (Klassen) $\{D^k, k = 1, 2, \dots, q\}$. Es werden dabei Verlustfunktionen $r_{kp}(\mathbf{X})$ definiert um den Verlust einer

Zuordnung $x \in D^p, (p = 1, 2, \dots, q)$ zu beschreiben, wenn x zur Population D^k gehört. Der Erwartungswert $E[r_{k,p}(x)]$ bezüglich der Verteilungsfunktion $f_k(x)$ in der Population p ist dann das Risiko, daß ein Patient aus der Population D^k der Population D^p zugeordnet wird. Ist die Entscheidung $x \in D^p$ getroffen, so wird das gesamte Risiko dieser Entscheidung durch

$$(4.5) \quad R_p = \sum_{k=1}^q E[r_{k,p}(x)]$$

bestimmt. Es ist weiter möglich, das durchschnittliche Risiko für alle Entscheidungen $x \in D^p, (p = 1, 2, \dots, q)$ zu erhalten

$$(4.6) \quad R = \sum_{p=1}^q P_p R_p, \quad ,$$

wenn die Einzelpopulationen $D^1, D^2, \dots, D^q$ im Verhältnis $P_1, P_2, \dots, P_q$ an der Gesamtpopulation beteiligt sind und die Gesamtpopulation aus der Summe der Einzelpopulationen zusammengesetzt ist. Die Verhältniszahlen $P_1, P_2, \dots, P_q$, die auch als apriori Wahrscheinlichkeiten bekannt sind, drücken Informationen über das Auftreten der Krankheiten (Populationen) aus. Das sind apriori Informationen also, solche Informationen, die vor dem laufenden Entscheidungsprozeß entstanden sind.

Die Minimierung der Fehlklassifikationsrate (4.6), unter der vereinfachenden Annahme, daß $r_{k,p} = 1$, (alle k, p) führt zu einer einfachen Entscheidungsregel. Entsprechend dieser Entscheidungsregel, die als Bayes-Strategie oder Likelihood-Quotienten-Kriterium bekannt ist (vergl. [2], Kap. 6), wird eine Realisation x der Population p zugeordnet, wenn

$$(4.7) \quad \frac{f_p(x)}{f_k(x)} > \frac{P_k}{P_p} \quad \text{für alle } k \neq p$$

4.3. Minimax - Lösung

Wenn keine apriori Wahrscheinlichkeiten vorliegen, ist die Bayes-Strategie nicht anwendbar. In solchen Fällen ist - im Sinne der Spieltheorie - der größtmögliche zu erwartende Verlust zu minimieren. Die ungünstigste apriori Verteilung der Populationen $D^1, D^2, \dots, D^q$ bezüglich der Gesamtpopulation ist die, bei der gilt

$$R_1 = R_2 = \dots = R_q \quad .$$

Das unter diesen Voraussetzungen zulässige Entscheidungsverfahren (Bayes-Strategie) ist als Minimax-Lösung bekannt. Es ist meistens der Fall, daß apriori Informationen über Krankheiten und Therapien vorliegen ; infolgedessen verwendet man die Minimax-Methode in der automatischen Diagnoseunterstützung selten.

4.4. Aposteriori Bayes-Wahrscheinlichkeiten

Läßt sich die gemeinsame Wahrscheinlichkeit $P(D^k, X \in D^k / X=x)$ der Population D^k und der Zugehörigkeit der Beobachtungsgröße x zur Population D^k als

$$(4.8) \quad P_k f_k(x)$$

schreiben, so kann die bedingte Wahrscheinlichkeit für $x \in D^k$ durch die Bayes-Formal ausgedrückt werden:

$$(4.9) \quad P(X \in D^k / X=x) = \frac{P_k f_k(x)}{\sum_{i=1}^q P_i f_i(x)}$$

Wenn man die Komponenten $x_1, x_2, \dots, x_m$ der Realisation $\mathbf{x}$ als voneinander unabhängig voraussetzt, kann die Bayes-Formel wie im folgenden umgeschrieben werden:

$$(4.10) \quad P(D^k/x_1, x_2, \dots, x_m) = \frac{P_k \prod_{i=1}^m f_k(x_i)}{\sum_{j=1}^q P_j \prod_{i=1}^m f_j(x_i)} .$$

Die Entscheidung fällt dann für die Population D^k wenn

$$P(D^k/x_1, x_2, \dots, x_m) > P(D^p/x_1, x_2, \dots, x_m) \quad \text{für alle } k \neq p.$$

Die Voraussetzung der statistischen Unabhängigkeit zwischen den Komponenten der Realisation $\mathbf{x}$ führt in den meisten Fällen zu einer Verfälschung der echten Situation. Sind die Merkmale stetig verteilt, so kann man dieses Problem umgehen, etwa dadurch, daß unter der Annahme einer multivariaten Normalverteilung die Abhängigkeiten mittels der Varianz-Kovarianz-Matrix dargestellt werden.

Für diskrete Merkmale gibt es mehrere Modelle. So können z.B. die polynomial Modelle [4][38] erwähnt werden, wobei die Verteilungsfunktion der als unabhängig betrachteten Merkmale mit einem Polynom multipliziert wird, das die gegenseitigen Abhängigkeiten der Merkmale enthält.

Die logistischen Modelle [14][23] drücken die Verteilungsfunktion der Realisation $\mathbf{x}$ durch Interaktionen verschiedener Ordnungen zwischen Variablen aus. Wie es auch aus der Theorie der mehrdimensionalen Kontingenztafeln [8][25] folgt, kann die Verteilungsfunktion als lineare Kombination von Haupteffekten und Interaktionen verschiedener Ordnung dargestellt werden.

Eine andere Art, diskrete Merkmale zu behandeln, ist ihr Verschmieren zu einer Normalverteilung. Danach können stetige Verfahren verwendet werden, mit der Einschränkung, daß hierbei viel der zur Trennung der Populationen beitragenden Information verlorengeht [12].

4.5. Klassifikation in normalverteilten Populationen

Wenn der Zufallsvektor $\mathbf{X}^T = [X_1, X_2, \dots, X_m]$ aus stetigen Komponenten zusammengesetzt ist, betrachtet man ihn in der Praxis oft als multivariat normalverteilt. D.h., daß die Zufallsvariable X_i mit einer gewissen Wahrscheinlichkeitsdichte jeden reellen Wert annehmen kann, wobei die größte Wahrscheinlichkeitsdichte dem Erwartungswert $\mu_i = E(X_i)$ entspricht. Die Varianz σ_i^2 beschreibt den erwarteten quadratischen Abstand zwischen dem Erwartungswert μ_i und der Zufallsvariable X_i

$$\sigma_i^2 = V(X_i) = E[(X_i - \mu_i)^2] \quad ,$$

wobei mit $E[\varphi(X)]$ der Erwartungswert der Funktion $\varphi(X)$ bezeichnet wurde. Für einen m-dimensionalen Zufallsvektor $\mathbf{X}^T = [X_1, X_2, \dots, X_m]$ ist dann der Mittelpunktvektor

$$(\boldsymbol{\mu}^k)^T = [\mu_1^k, \mu_2^k, \dots, \mu_m^k]$$

bezüglich der Population D^k . Der Varianz entspricht, in diesem Fall, die Varianz-Kovarianz-Matrix $\boldsymbol{\Sigma}^k = \{\sigma_{ij}^k\}$, wobei $\sigma_{ii}^k = (\sigma_i^k)^2 = V(X_i)$ und $\sigma_{ij}^k = \text{Cov}(X_i, X_j)$ in der Population D^k sind. Die normale Dichtefunktion an der Stelle $\mathbf{x}$ schreibt sich dann

$$(4.11) \quad f_k(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}^k|^{1/2}} \exp\left[(-1/2)(\mathbf{x} - \boldsymbol{\mu}^k)^T (\boldsymbol{\Sigma}^k)^{-1} (\mathbf{x} - \boldsymbol{\mu}^k)\right]$$

Betrachten wir den Fall nur zweier normalverteilten Populationen

$$(4.12) \quad \begin{aligned} f_1(\mathbf{x}) &: N(\boldsymbol{\mu}^1, \boldsymbol{\Sigma}^1) \\ f_2(\mathbf{x}) &: N(\boldsymbol{\mu}^2, \boldsymbol{\Sigma}^2) \quad , \end{aligned}$$

so gibt es eine orthogonale Transformation, die die beiden

Matrizen Σ^1 und Σ^2 auf Diagonalform bringt und den Mittelpunktvektor einer der Populationen in den Koordinatenursprung verschiebt. Nach [24] lautet eine solche Transformation

$$(4.13) \quad \mathbf{y} = \mathbf{A}^T (\Sigma^1)^{-1/2} (\mathbf{x} - \mu^1) \quad ,$$

wobei $\mathbf{A}$ eine orthogonale Matrix ist, so daß

$$\mathbf{A}^T (\Sigma^1)^{-1/2} \Sigma^2 [(\Sigma^1)^{-1/2}]^T \mathbf{A} = \mathbf{D} \quad .$$

Im weiteren wird mit $\mathbf{D}$ folgende Diagonalmatrix bezeichnet:

$$\mathbf{D} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_m^2 \end{bmatrix}$$

Nach der Transformation (4.13) genügt der neue Beobachtungsvektor $\mathbf{y}$ in der Population D^1 einer Normalverteilung $N(\mathbf{0}, \mathbf{I})$ und in der Population D^2 einer Normalverteilung $N(\mu^*, \mathbf{D})$, wobei

$$(4.14) \quad \mu^* = \mathbf{A}^T (\Sigma^1)^{-1/2} (\mu^2 - \mu^1) \quad .$$

Es ist hier wichtig zu bemerken, daß die Fläche der konstanten Wahrscheinlichkeitsdichte für (4.11) ein m -dimensionales Hyperellipsoid beschreibt. Demzufolge wird die Fläche der der konstanten Wahrscheinlichkeit nach der Transformation (4.13) in der Population D^1 eine Hyperkugel der Form

$$(4.15) \quad \sum_{i=1}^m y_i^2 = K_1$$

und in der Population D^2 ein Hyperellipsoid der Form

$$(4.16) \quad \sum_{i=1}^m \frac{(y_i - \mu_i^*)^2}{\sigma_i^2} = K_2 \quad .$$

Daraus ist leicht zu erkennen, daß die Schnittstelle zwischen

den beiden Ellipsoiden (falls sie existiert) eine Hyperebene der Gestalt

$$(4.17) \quad L(\mathbf{y}) = \mathbf{y}^T \mathbf{a} + b = 0$$

ist, wobei $\mathbf{a}$ einen Koeffizientenvektor und b eine Konstante darstellt.

Die Hyperellipsoide (4.15) und (4.16) schließen also im m -dimensionalen Raum zwei Zonen ein, wo eine Beobachtungsgröße $\mathbf{x}$ mit einer Wahrscheinlichkeitsdichte $f_1(\mathbf{x}) > e^{k_1}$ zu der Population D^1 gehört bzw. mit einer Wahrscheinlichkeit $f_2(\mathbf{x}) > e^{k_2}$ zu der Population D^2 gehört. Die Pegel der Wahrscheinlichkeiten f_1 und f_2 können so ausgewählt werden, daß eine Schnittstelle der Form (4.17) immer existiert. Eine solche Hyperebene teilt den Entscheidungsraum in zwei Halbräume ein (Abb.3).

Wenn der Punkt $\mathbf{x}$ in die Schnittmenge der beiden Gebiete fällt, prüft man, ob er sich auf der einen Seite ($L(\mathbf{x}) < 0$) oder auf der anderen Seite ($L(\mathbf{x}) > 0$) der Ebene befindet, um seine Zugehörigkeit zu einer der Populationen zu bestimmen.

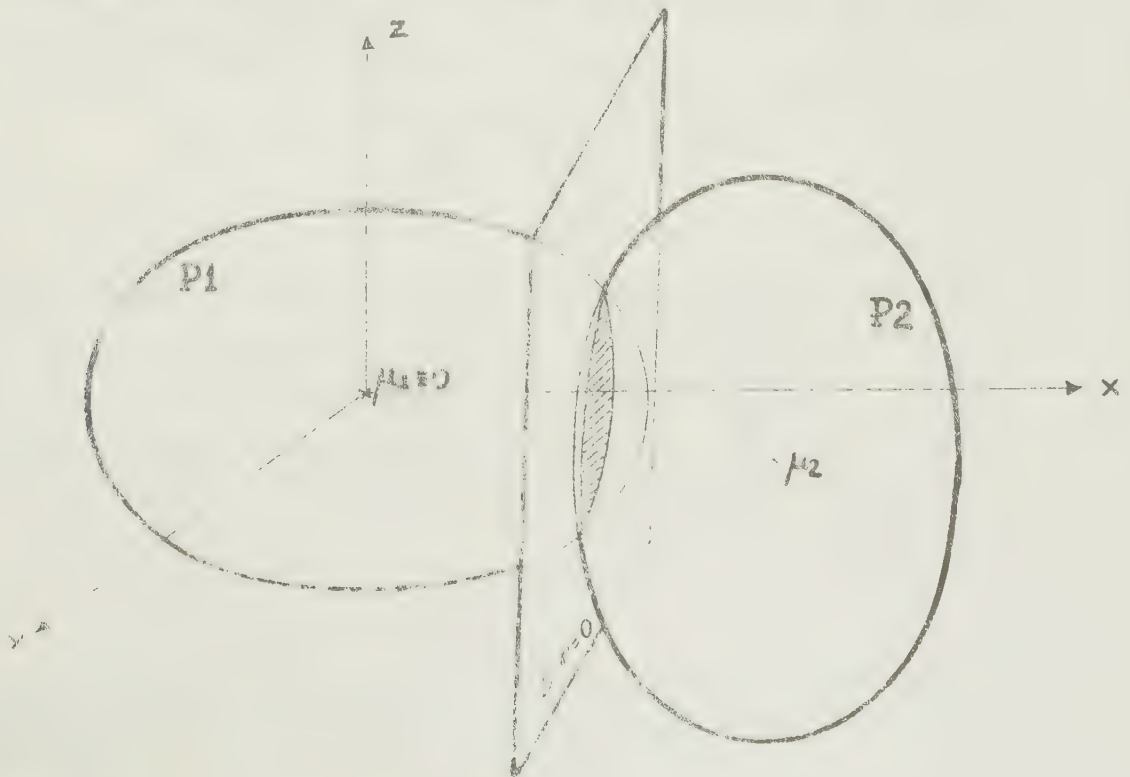


Abb.3 - Räumliche Darstellung der Diskriminanzfunktion

5. Fehler beim Abschätzen und Klassifizieren mit vielen Variablen

Es wurden im vorangegangenen Kapitel verschiedene wahrscheinlichkeitstheoretische Modelle beschrieben, die eine Dichtefunktion $f(\mathbf{x}, \boldsymbol{\theta})$ in Abhängigkeit von der Beobachtungsgröße $\mathbf{x}$ und vom Parametervektor $\boldsymbol{\theta}$ (Für die Normalverteilung gilt $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}\}$) voraussetzen. Die Klassifikationsregeln wurden unter der Voraussetzung aufgestellt, daß der Parametervektor $\boldsymbol{\theta}$ bekannt ist. In Wirklichkeit muß er aber von der verfügbaren Stichprobe geschätzt werden.

Infolgedessen wird vermutet, daß eine Stichprobe $\mathbf{Z} = \{\mathbf{x}^i, (i = 1, 2, \dots, N)\}$ von N unabhängigen Zufallsvektoren aus einer statistischen Population zur Verfügung steht. Unter der Voraussetzung der statistischen Unabhängigkeit zwischen den Probanden $\mathbf{x}^i$ ergibt sich die Möglichkeit, ihre gemeinsame Dichtefunktion durch Multiplikation der Randdichtefunktionen abzuleiten:

$$(5.1) \quad f(\mathbf{Z}, \boldsymbol{\theta}) = \prod_{i=1}^N f(\mathbf{x}^i, \boldsymbol{\theta})$$

Betrachtet man jetzt den Parametervektor $\boldsymbol{\theta}$ als Realisation eines Zufallsvektors $\boldsymbol{\Theta}$, so kann die Funktion (5.1) als eine gemeinsame Verteilungsfunktion der Zufallsgrößen $\mathbf{Z}$ und $\boldsymbol{\Theta}$ betrachtet werden. Für eine gewisse Stichprobe $\mathbf{z}$ hängt dann die Funktion

$$(5.2) \quad f(\mathbf{z}, \boldsymbol{\Theta})$$

nur vom Zufallsvektor $\boldsymbol{\Theta}$ ab. Sie ist als Likelihood-Funktion der Parameter bekannt und es ist zu erwarten,

daß für eine vorgegebene Stichprobe $\mathbf{z}$ die beste Schätzung des Parametervektors Θ diese Funktion maximieren wird. Daraus folgt sofort die Gleichung

$$(5.3) \quad \frac{\partial f(\mathbf{z}, \Theta)}{\partial \Theta} = 0 ,$$

welche die sogenannte Maximum-Likelihood-Schätzung $\hat{\Theta}_{ML}$ der Parameter liefert. Wird diese Schätzung in den Klassifikationsregeln vom vorangegangenen Kapitel eingesetzt, so erhält man ein estimatives Diagnosemodell. Die estimativen Diagnosemodelle führen in der Regel zu optimistischen Ergebnissen, die sich von der Realität entfernen[1].

Aufgrund der abgeschätzten Parameter kann die aposteriori Dichtefunktion ¹⁾ wie im folgenden geschrieben werden

$$(5.4) \quad f(\mathbf{x} \in D^k / \mathbf{x}, \mathbf{z}) \propto P_k f(\mathbf{x} / D^k, \hat{\Theta}_{ML}(\mathbf{z})) ,$$

wobei durch $\hat{\Theta}_{ML}(\mathbf{z})$ die Abhängigkeit des Parametervektors von der Stichprobe $\mathbf{z}$ und mit $\propto$ die Proportionalität ausgedrückt wird.

Geht man von der Bayes'schen Schätzungsmethode aus, so tritt tritt ein neuer Patient, dessen Merkmalsvektor $\mathbf{x}$ ist, mit folgender Wahrscheinlichkeit auf

$$(5.5) \quad f(\mathbf{x} / D^k, \mathbf{z}) = \int_{S_{\Theta}} f(\mathbf{x} / D^k, \Theta) f(\Theta / \mathbf{z}) d\Theta ,$$

wobei S_{Θ} den Parameterraum darstellt. Die aposteriori Wahrscheinlichkeitsdichte wird in diesem Fall

$$(5.6) \quad f(\mathbf{x} \in D^k / \mathbf{x}, \mathbf{z}) \propto P_k f(\mathbf{x} / D^k, \mathbf{z}) ,$$

wobei $f(\mathbf{x} / D^k, \mathbf{z})$ als prädiktive Dichtefunktion bekannt ist. Dieses Vorgehen führt zu den prädiktiven Diagnosemodellen [1], ([2], S. 385). Es ist hier deutlich zu erkennen, daß die prädiktive Dichtefunktion (5.5) als Durchschnittswert über alle Parameter mit Hilfe der aposteriori Dichtefunktion $f(\Theta / \mathbf{z})$ berechnet wurde. Unter Anwendung der

1) Siehe Kapitel 4. , Formel (4.9)

Bayes-Regel kann $f(\theta/z)$ von der apriori Dichtefunktion der Parameter $p(\theta)$ und der Likelihood-Funktion der geschätzten Parameter $f[\hat{\theta}(z)/D^k, \theta]$ (Diese Funktion beschreibt die Verteilung der Schätzungen $\hat{\theta}(z)$ unter der Bedingung, daß die wahren Parameter θ in der Population D^k bekannt sind) wie im folgenden bestimmt werden:

$$(5.7) \quad f(\theta/z) \propto f[\hat{\theta}(z)/D^k, \theta] p(\theta) \quad .$$

Der Anhang A enthält ein Beispiel für normalverteilte Daten, wobei diese Prozedur in Einzelheiten dargestellt ist.

5.1. Die Fehlklassifikationsrate als Funktion der Anzahl der Dimensionen

Es ist bekannt([2], S.142), daß aufgrund des Quotienten $l(x) = f_1(x)/f_2(x)$ eine vollständige Klasse von Klassifikationsregeln wie im folgenden aufgebaut werden kann

$$\xi = \begin{cases} 0 & \text{wenn } l(x) > c, \quad x \text{ gehört zu } D^1 \\ 1 & \text{wenn } l(x) < c, \quad x \text{ gehört zu } D^2 \\ 0 \text{ mit W. } \gamma & , \quad x \text{ gehört zu } D^1 \text{ mit der} \\ 1 \text{ mit W. } 1-\gamma & \text{Wahrscheinlichkeit } \gamma \text{ und zu } D^2 \\ & \text{mit der Wahrscheinlichkeit } 1-\gamma \end{cases}$$

Für verschiedene c und γ bilden die Klassifikationsregeln ξ eine vollständige Klasse, wobei für $\gamma=0$ die nicht zufälligen Klassifikationsregeln bestimmt werden. Im folgenden wird mit $\alpha_i(\xi)$, ($i=1,2$) die Wahrscheinlichkeit bezeichnet, daß ein aus der Population D^i zufällig ausgewähltes Element falsch klassifiziert wird.

In Wirklichkeit werden die Parameter der Likelihood-Funktionen $f_1(x) = f(x, \theta_1)$ und $f_2(x) = f(x, \theta_2)$ aus einer Stichprobe abgeschätzt. Da ξ , die auf

$l(x) = f(x, \theta_1) / f(x, \theta_2)$ basierende Klassifikationsregel darstellt, wird mit ξ^* diejenige Regel bezeichnet, die sich auf dem geschätzten Likelihood-Quotienten $\hat{l}(x) = f(x, \hat{\theta}_1) / f(x, \hat{\theta}_2)$ stützt. Da ξ^* aufgrund einer Stichprobe mit N_1 Individuen aus der Population D^1 und N_2 Individuen aus der Population D^2 abgeleitet wurde, ist es als eine Zufallsvariable mit den möglichen Werten 0, γ oder 1 für jeden x zu betrachten. Wenn nur ein Individuum, z.B. $X = x_0$, zu klassifizieren ist, wird nach der Wahrscheinlichkeit $\Pr[\xi^*(x_0) = \xi(x_0)]$ gesucht. Wird eine größere Anzahl von Individuen klassifiziert, so sind folgende Größen von Interesse:

$$(5.8) \quad \alpha_1(\xi^*) = E_1[\xi^*(X)]$$

$$(5.9) \quad E[\alpha_1(\xi^*)] = E[E_1(\xi^*(X))] = \Pr[\xi^*(X) = 1/D^1] \quad ,$$

wobei E_1 und E Operatoren zur Berechnung des Erwartungswertes bezüglich der Dichte $f(x, \theta_1)$ bzw. der gemeinsamen Sample-Dichtefunktion der Vektoren $\hat{\theta}_1$ und $\hat{\theta}_2$ darstellen. Hierbei beschreibt (5.8) die bedingte Wahrscheinlichkeit, daß bei Vorgabe von ξ^* , ein zufällig ausgewähltes Individuum aus D^1 nicht richtig klassifiziert wird. Dagegen beschreibt (5.9) die unbedingte Wahrscheinlichkeit, daß ein zufällig ausgewähltes Stichproben-Paar von D^1 und D^2 zu einer Klassifikationsregel ξ^* führt, die ein zufälliges Individuum aus D^1 nicht richtig klassifizieren wird. Die letzte Größe hilft zunächst eine optimale Klassifikationsregel ξ^* zu finden. Danach liefert (5.8) die Fehlklassifikationsrate für D^1 unter der Bedingung, daß die Klassifikationsregel ξ^* verwendet wurde. Geht man von der linearen Diskriminanzfunktion (4.18) aus, so erhält man unter der Vereinfachung $\ln(P_1/P_2) = 0$ für den Likelihood-Quotienten

$$(5.10) \quad l(x) = (\mu^1 - \mu^2)^T \Sigma^{-1} [x - (1/2)(\mu^1 + \mu^2)] \quad .$$

Die auf diesem Kriterium basierende Klassifikationsregel wird im folgenden mit ξ bezeichnet. Man bezeichnet weiter mit $G(t)$ die Funktion

$$G(t) = \int_{-\infty}^t g(y) dy, \quad ,$$

wobei

$$g(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} y^2\right) .$$

Unter der Annahme einer Normalverteilung wird die Fehlklassifikationsrate für ξ in D^1 durch

$$\alpha_1(\xi) = G\left(-\frac{1}{2} d\right)$$

ausgedrückt [28], wobei d die positive Wurzel des Mahalanobis- Abstandes $d^2 = (\mu^1 - \mu^2)^T \Sigma^{-1} (\mu^1 - \mu^2)$ darstellt.

Unter der Annahme eines estimativen Diagnosemodells (s.S.41) werden Maximum- Likelihood-Schätzungen (s.Anhang A) $\bar{x}^1, \bar{x}^2, \hat{S}$ der Parameter μ^1, μ^2 bzw. Σ in den Quotienten (5.10) eingesetzt. Der geschätzte Likelihood-Quotient wird dann:

$$(5.11) \quad \hat{l}(x) = (\bar{x}^1 - \bar{x}^2)^T \hat{S}^{-1} \left[x - \frac{1}{2} (\bar{x}^1 + \bar{x}^2) \right] .$$

Wenn ξ^* die aufgrund von $\hat{l}(x)$ aufgebaute Klassifikationsregel ist, dann erhält man für $\ln(P_1/P_2)=0$ folgende Fehlklassifikationsrate:

$$(5.12) \quad \alpha_1(\xi^*) = G \left[\frac{-(\mu^1)^T \hat{S}^{-1} (\bar{x}^1 - \bar{x}^2) + (1/2)(\bar{x}^1 + \bar{x}^2)^T \hat{S}^{-1} (\bar{x}^1 - \bar{x}^2)}{\sqrt{[(\bar{x}^1 - \bar{x}^2)^T \hat{S}^{-1} \Sigma \hat{S}^{-1} (\bar{x}^1 - \bar{x}^2)]}} \right] .$$

$\bar{x}^1 \neq \bar{x}^2$ wird selbstverständlich dabei vorausgesetzt. Es ist deutlich zu sehen, daß $\alpha_1(\xi^*)$ von den Schätzungen $\bar{x}^1, \bar{x}^2, \hat{S}^{-1}$ sowie von den unbekannten Parametern μ^1 und Σ abhängig ist. Die Eigenschaften der Verteilung (5.12) können aber geschätzt werden, wenn statt μ^1 und Σ die entsprechenden geschätzten Parameter eingesetzt werden.

Man kann daraus ein angenähertes Wahrscheinlichkeitsintervall für $\alpha_1(\xi^*)$ bestimmen.

Es ist von Interesse, die Funktion $E[\alpha_1(\xi^*)]$ zu bestimmen. Sie hängt nur von den Größen d , N_1 , N_2 und m ab. Da man $E[\alpha_1(\xi^*)]$ nicht genau tabellieren konnte [wie z.B. (5.12)], ist in diesem Fall die asymptotische Entwicklung von Okamoto [40] zu verwenden. Gegeben wurde diese Entwicklung bis zu Termen des Grades $N_1^{-2}, N_2^{-2}, (N_1 N_2)^{-1}, (N_1 n)^{-1}, (N_2 n)^{-1}, n^{-2}$ mit $n = N_1 + N_2 - 2$. In der Tabelle 1 sind Werte der Funktion $E[\alpha_1(\xi^*)]$ für verschiedene D-Werte und verschiedene Werte der Dimension m angegeben. Daraus ist zu ersehen, daß die Fehlklassifikationsrate mit der Anzahl m der Dimensionen steigt. Dieses ist aus dem Zusammenhang zwischen den abzuschätzenden Parametern und der Anzahl der Variablen leicht zu erklären. Für normalverteilte Daten steigt die Zahl der abzuschätzenden Parameter sehr stark mit der Anzahl der Variablen m , nämlich im Verhältnis $m(m+3)/2$. Für große Dimensionen reichen dann die verfügbaren Stichproben für das Abschätzen der Parameter nicht aus.

Unter der Annahme eines prädiktiven Diagnosemodells, kann die lineare Diskriminanzfunktion nicht mehr verwendet werden. Der Likelihood-Quotient läßt sich unter dieser Voraussetzung folgendermaßen schreiben ([41], S. 576) :

$$l(\mathbf{x}) = L_{12} \frac{\left[1 + \frac{N_1}{N_1^2 - 1} (\mathbf{x} - \bar{\mathbf{x}}^1)^T (\mathbf{S}^1)^{-1} (\mathbf{x} - \bar{\mathbf{x}}^1) \right]^{N_1/2}}{\left[1 + \frac{N_2}{N_2^2 - 1} (\mathbf{x} - \bar{\mathbf{x}}^2)^T (\mathbf{S}^2)^{-1} (\mathbf{x} - \bar{\mathbf{x}}^2) \right]^{N_2/2}}$$

wobei die Konstante L_{12} folgende Form hat:

$$L_{12} = \frac{P_1}{P_2} \left[\frac{|(N_1 - 1)\mathbf{S}^1|}{|(N_2 - 1)\mathbf{S}^2|} \right]^{1/2} \frac{\Gamma\left(\frac{N_2}{2}\right) \Gamma\left(\frac{N_1 - m}{2}\right)}{\Gamma\left(\frac{N_1}{2}\right) \Gamma\left(\frac{N_2 - m}{2}\right)} \left[\frac{N_2 (N_1 + 1)}{N_1 (N_2 + 1)} \right]^{m/2}$$

Table 1 (Werte der Funktion $E[\alpha_1(\xi^*)]$, berechnet nach der asymptotischen Entw. von M. Okamoto[40]

$N_1 = N_2 = 50 \quad , \quad n = 98$			$N_1 = N_2 = 100 \quad , \quad n = 198$			$N_1 = N_2 = 1000 \quad , \quad n = 1998$			
$D = 4$	$D = 4$	$D = 8$	$D = 1$	$D = 4$	$D = 8$	$D = 1$	$D = 4$	$D = 8$	
1	0.55541	0.02329	0.043443	0.30900	0.02302	0.043303	0.30897	0.02278	0.043180
2	0.31384	0.02394	0.043757	0.31120	0.02334	0.043453	0.30880	0.02281	0.043195
3	0.31804	0.02462	0.044100	0.31330	0.02369	0.043609	0.30902	0.02284	0.043209
5	0.32530	0.02601	0.044862	0.31750	0.02436	0.043941	0.30946	0.02291	0.043238
7	0.33276	0.02748	0.045728	0.32140	0.02508	0.044298	0.30990	0.02298	0.043267
10	0.34111	0.03094	0.047225	0.32690	0.02617	0.044884	0.31055	0.02308	0.043311
20	0.35583	0.03801	0.051271	0.34170	0.03010	0.046028	0.31957	0.02342	0.043463
50	0.37025	0.04711	0.05779	0.35590	0.03516	0.047023	0.32594	0.02390	0.043575

Hierbei wurde die Bezeichnung $\Gamma(x)$ für die Gamma-Funktion verwendet. Für eine solche komplizierte Klassifikationsregel war es bis jetzt unmöglich, eine Verteilung der Fehlklassifikationsrate $\alpha_1(\xi^*)$ oder $E[\alpha_1(\xi^*)]$ zu finden. Hierfür können aber Monte-Carlo-Zufallsexperimente eingesetzt werden, wobei mit Hilfe von simulierten Zufallsdaten eine Schätzung der Fehlklassifikationsrate möglich ist. Für die quadratische Diskriminanzanalyse liegen

bereits praktische Ergebnisse vor [12] (Abb.4).

Aus der Abbildung 4 kann man schließen, daß je mehr Variablen bei der Abschätzung der Parameter berücksichtigt werden, die Fehlklassifikationsrate desto größer sein wird.

5.2. Systematische Fehler und redundante Variablen

Systematische Fehler treten meistens bei der Verwendung von falschen apriori Informationen oder redundanten Variablen auf, wobei als redundant diejenige Variable bezeichnet wird, die keine, oder eine geringe Information zur Trennung der Populationen enthält [31]. Daß redundante Variablen systematische Fehler verursachen, ist auch durch die Monte-Carlo-Experimente bewiesen worden ¹⁾. Aus der Abbildung 4 ist zu ersehen, daß die Trennfähigkeit des quadratischen Trennalgorithmus bei Hinzufügen von nicht zur Trennung beitragenden Variablen sinkt, da dadurch ein Teil der trennenden Information zum Abschätzen der Parameter verlorengeht. Bemerkenswert ist in der Tabelle 1 die Abhängigkeit der Fehlklassifikationsrate von der Anzahl der Variablen m . Um diese wichtigen Funktionen besser untersuchen zu können, wurden einige davon in der Abbildung 5 dargestellt.

1) Siehe Bericht MD 95 des Instituts für Medizinische Datenverarbeitung der GSF München

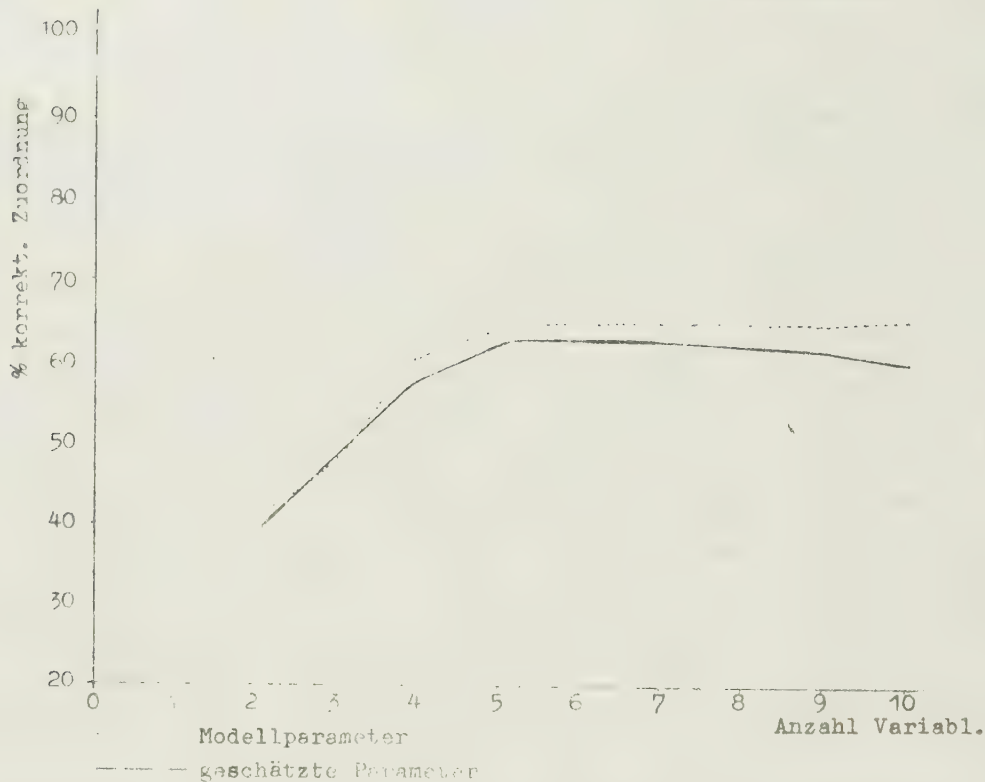


Abb. 4 - Neuordnung mittels quadratischer Diskriminanzanalyse
(aus dem Bericht MD 95 des IMD - GSF - München[12]).

Für die verwendeten Stichprobenmengen kann man aus der Abbildung 5 eindeutig schließen, daß die Fehlklassifikationsrate stark mit dem Abstand D sinkt. Wenn aber 50 Variablen zur Trennung der Populationen verwendet werden, steigt die Fehlklassifikationsrate für $D=1$ oder für $D=8$ um das 1,15- fache, bzw. um das 5,52- fache im Vergleich zu dem Fall, daß nur eine Variable in die Berechnung einbezogen wurde. Setzt man die Definition der Trenninfor- mation vom Kapitel 7 als bekannt voraus, so kann man sagen, daß je mehr Trenninformation vorhanden ist (D größer ist), desto mehr davon durch eine Erhöhung der Dimension anschließend verloren geht. Für den Fall, daß kleine Stichproben- mengen vorliegen und der Abstand D gering oder gleich Null ist, kann die Fehlklassifikationsrate mit der Anzahl der Dimensionen sinken ¹⁾. Dieses paradoxe Ergebnis ist dadurch zu erklären, daß, wenn die Stichprobenmengen eine gewisse Anzahl unterschreitet (ungefähr gleich der Anzahl der Variablen), und dadurch die gesamte Trenninformation

1) Siehe Tabelle 1 ($n = n = 50$ und $m = 50$)

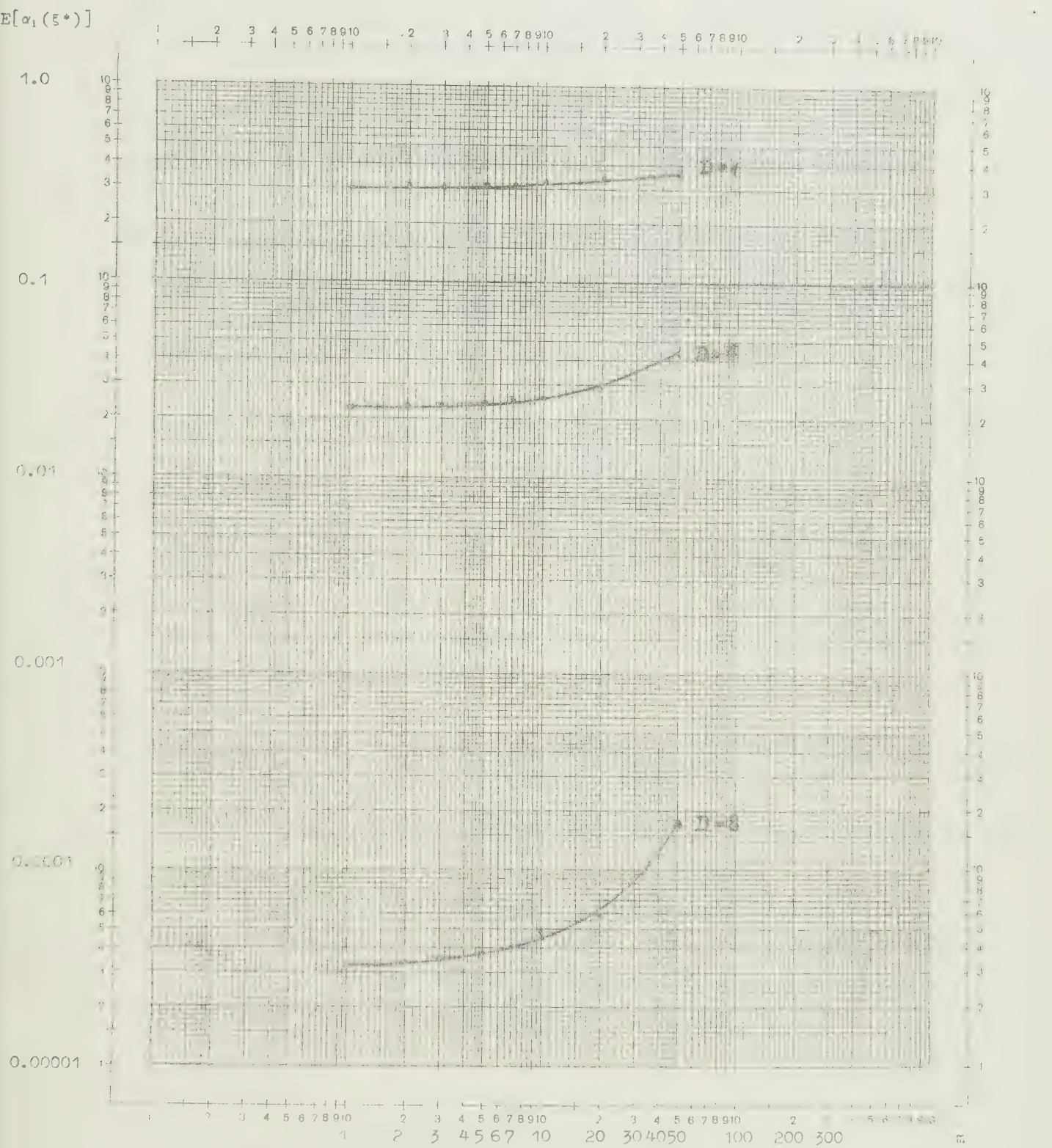


Abb. 5 - Fehlklassifikationsrate als Funktion der Anzahl der Dimensionen
 für $N_1 = N_0 = 100$ und $n = 198$
 (Punkte nach Okamoto[40] siehe Tabelle 1)

sehr gering ist, eine Erhöhung der Dimension mehr Information über die Population liefert als vernichtet.

Für große Dimensionen ist es zweifellos möglich, die Fehlklassifikationsrate durch Selektion der nichtredundanten Variablen zu verringern. Ein Maß der in jeder Variablen enthaltenen Trenninformation ist dafür erforderlich. Man kann danach auch ein Kriterium finden, um redundante von nichtredundanten Variablen trennen zu können.

6. Bekannte Verfahren zur Reduktion der Dimensionen

Da normalverteilte Populationen einen günstigen mathematischen Formalismus anbieten, sind fast alle herkömmlichen Verfahren zur Reduktion der Dimensionen unter Voraussetzung der Normalität entwickelt worden.

Werden drei normalverteilte Populationen als Beispiel angenommen, so können die Varianz-Kovarianz-Matrizen mit Σ^k , ($k = 1, 2, 3$) und die entsprechenden Mittelpunktsvektoren mit μ^k , ($k = 1, 2, 3$) bezeichnet werden. Man kann weiter voraussetzen, daß zur Parameterschätzung insgesamt N Probanden, eingeteilt in N_1, N_2, N_3 , für die jeweilige Population zur Verfügung stehen. Die nach der Maximum-Likelihood-Methode geschätzten Parameter schreiben sich

$$(6.1) \quad \bar{x}^k = \frac{1}{N_k} \sum_{i=1}^{N_k} x^i, \quad S^k = \frac{1}{N_k - 1} \sum_{i=1}^{N_k} (x^i - \bar{x}^k)(x^i - \bar{x}^k)^T, k=1, 2,$$

$$(6.2) \quad \bar{x} = \frac{1}{N} \sum_{k=1}^3 \sum_{i=1}^{N_k} x^i, \quad S = \frac{1}{N - 1} \sum_{k=1}^3 \sum_{i=1}^{N_k} (x^i - \bar{x}^k)(x^i - \bar{x}^k)^T,$$

wobei $\bar{x}^k, S^k$ den abgeschätzten Mittelpunktsvektor bzw. Varianz-Kovarianz-Matrix der Population D^k darstellt und mit $\bar{x}, S$ das abgeschätzte Hauptzentrum bzw. die gepoolte Varianz-Kovarianz-Matrix bezeichnet wurde.

6.1. Hauptkomponentenmethode

In diesem Fall wird die gepoolte Varianz-Kovarianz-Matrix S oder die entsprechende Korrelationsmatrix untersucht. Die Probanden x^i werden einer Transformation

$$(6.3) \quad y^i = A x^i, \quad ,$$

zum Maximieren der Varianzen der neuen Variablen unterworfen. Die j -te Variable des Vektors $\mathbf{y}$ schreibt sich dann

$$y_j = \alpha_j^T \mathbf{x} \quad ,$$

wobei α_j die j -te Spalte der Transformationsmatrix $\mathbf{A}$ darstellt. Fängt man mit y_1 an, so muß die Varianz

$$\sigma^2_1 = \alpha_1^T \mathbf{S} \alpha_1$$

maximiert werden. Mit Hilfe eines Lagrangeschen Multiplikators und durch das Ableiten ([41], S.286) erhält man unter der Bedingung $\alpha_1^T \alpha_1 = 1$ folgende Gleichung

$$(6.4) \quad (\mathbf{S} - \lambda_1 \mathbf{I}) \alpha_1 = \mathbf{0} \quad .$$

Die Lösung folgt aus der Determinantengleichung

$$(6.5) \quad |\mathbf{S} - \lambda \mathbf{I}| = 0 \quad ,$$

wobei λ die Unbekannte ist. Nimmt man reelle und positive Lösungen für die Gleichung (6.5) an, so kann dieses einfache Eigenwertproblem gelöst werden. Die größte Lösung λ_1 der Gleichung (6.5) ist dann der erste Eigenwert der Matrix $\mathbf{S}$ und α_1 stellt den entsprechenden normierten Eigenvektor dar. Der Eigenwert λ_1 ist auch als die Varianz der neuen Variablen y_1 ($\sigma^2_1 = \lambda_1$) zu interpretieren ([41], S.286). Wird nach einer neuen Variablen y_2 gesucht, die orthogonal zu y_1 ist und deren Varianz

$$\sigma^2_2 = \alpha_2^T \mathbf{S} \alpha_2$$

ein Maximum erreichen muß, so erhält man

$$y_2 = \alpha_2^T \mathbf{x} \quad ,$$

wobei α_2 den zum nächsten Eigenwert λ_2 zugehörigen Eigenvektor darstellt. Am Ende dieser Prozedur erhält man ebenfalls einen m -dimensionalen Variablenvektor $\mathbf{y}$ dessen

Komponenten voneinander unabhängig oder nicht korreliert sind, je nachdem die Varianz-Kovarianz-Matrix S bzw. die entsprechende Korrelationsmatrix in die Berechnung einbezogen wurde.

Der Raum der neuen Variablen y_i ist als Hauptachsenraum und die Transformation (6.3) als Hauptachsentransformation bekannt. Da in diesem Raum die Variablen voneinander unabhängig bzw. nicht korreliert sind, können einige Variablen davon vernachlässigt werden ohne andere Variablen zu beeinflussen. Die Größe der Variablenvarianz ist als Auswahlkriterium der Variablen zu verwenden. Das führt zur Selektion der Variablen mit der größten Varianz. Solche Variablen, die ungefähr als Konstanten auftreten, werden dementsprechend als unwichtig für den Trennprozeß betrachtet.

6.2. Methode des mehrfachen Korrelationskoeffizienten

Diese Methode geht davon aus, die Korrelation zwischen einer Variablen und allen anderen von der vorgegebenen Stichprobe zu berechnen. Sie ist eine sequentielle Methode, da die Variablen nacheinander in die Berechnung einbezogen werden. Es wird diejenige Variable ausgewählt, die den größten Beitrag für den mehrfachen Korrelationskoeffizienten liefert.

Unter der Berücksichtigung, daß der mehrfache Korrelationskoeffizient eine Funktion der paarweisen Varianzen ist, kann man diese Methode prinzipiell mit der Hauptachsentransformation gleichstellen.

Verschiedene Varianten dieser Methode führen nicht immer zum selben Ergebnis; ist das jedoch der Fall, so ist das erreichte Ergebnis nicht optimal ([31], S.98).

6.3. Wilks'- Lambda - Test

Die Partition der Total-Matrix in Zwischen-Populationen-Matrix und Innerhalb-der-Populationen-Matrix (vgl. [13], Kap. 8)

$$T = A + \Sigma$$

kann im Falle von q Populationen zu einem Test für die folgende Hypothese (Nullhypothese)

$$H_0 : \mu^1 = \mu^2 = \dots = \mu^q = \mu$$

führen (s. Anhang B). 1932 hat Wilks [5] als Kriterium für die Annahme oder Zurückweisung der Hypothese H_0 folgende Statistik erfunden

$$(6.6) \quad \Lambda = \frac{|\Sigma|}{|T|} \quad .$$

Es ist einsichtig, daß je mehr dieser Quotient - auch als Wilks'-Lambda-Quotient bekannt - zur Einheit strebt, um so weniger Unsicherheit mit der Annahme der Hypothese H_0 verbunden sein wird.

Unter der Voraussetzung, daß der Test auf gleichen Mittelpunktsvektoren¹⁾ ein Maß der Diskriminanzleistung für den verwendeten Variablensatz darstellt, wird im folgenden versucht, den Λ - Test für einzelne Variablen durchzuführen. Dafür wird zunächst erwähnt, daß die Verteilung des Lambda-Quotienten mit Hilfe von Rao's F-Approximation (F-Verteilung) berechnet werden kann ([42], S. 261).

1) Die Schlußbemerkung enthält eine ausführliche Diskussion dieser Voraussetzung.

Um diese F-Approximation zu beschreiben sind folgende Hilfsgrößen erforderlich

$$s = \sqrt{\left\{ \frac{m^2(q-1)^2 - 4}{m^2 + (q-1)^2 - 5} \right\}}$$

$$n_1 = m(q-1)$$

$$n_2 = s \left[(N-1) - \frac{m + (q-1) + 1}{2} \right] - \frac{m(q-1) - 2}{2},$$

wobei m die Anzahl der Variablen, q die Anzahl der Populationen und N die Anzahl aller Probanden darstellen. Die F-Approximation lautet dann für die Größe $\rho = \Lambda^{1/s}$

$$(6.7) \quad F_{n_2}^{n_1} = \left(\frac{1-\rho}{\rho} \right) \left(\frac{n_1}{n_2} \right),$$

wobei der Zähler des Varianz-Bruches n_1 Freiheitsgrade und der Nenner n_2 Freiheitsgrade hat ([13], Kap.8).

Schließlich ist hier noch zu bemerken, daß die Diagonalelemente der Matrizen $\mathbf{A}$ und $\mathbf{\Sigma}$ Zwischen-Populationen- bzw. Innerhalb-der-Populationen-mittlere-Quadrate für einzelne Variablen liefern. Werden damit Tests für einzelne Variablen konstruiert, so ist es möglich sie nur in Abhängigkeit von der Hypothese H_0 zu verwenden. Wenn die Hypothese sich als falsch erwiesen hat, können solche univariaten Tests eingesetzt werden. Man kann dadurch einen Eindruck über die Diskriminanzleistung einzelner Variablen gewinnen.

Unabhängige Tests für univariaten Hypothesen wurden in der Arbeit von D.F.Morrison ([39], Kap.5) diskutiert. Die Analyse erweist sich als sehr komplex, wobei die Rechenmethoden und dadurch auch die Ergebnisse nicht leicht zu interpretieren sind.

6.4. Reduktion des Diskriminanzraumes

In der Diskriminanzanalyse wird nach einer linearen Transformation

$$(6.8) \quad y = \Gamma x$$

gesucht [31], wobei der Quotient aus der Zwischen-Populationen-Varianzen und der Innerhalb-der-Populationen-Varianzen maximiert werden muß.

Die Matrix Γ bestimmt die Transformation des ursprünglichen Stichprobenraumes in einen Diskriminanzraum.

Es ist weiter möglich den Quotienten Λ , der im Paragraphen 6.3. beschrieben wurde, als Funktion der Eigenwerten der Matrix $\Sigma^{-1}A$ wie im folgenden darzustellen ([13], Kap.9)

$$(6.9) \quad \Lambda = \prod_{k=1}^q \frac{1}{1 + \lambda_k}$$

Wird auf p von q Diskriminanzfunktionen verzichtet, so lautet der Chi-Quadrat-Signifikanztest für die verlorengegangene Diskriminanzleistung ([13], S.249)

$$\chi^2 = - \left(n - \frac{m+q}{2} - 1 \right) \ln \Lambda'$$

mit $(m-p)(q-p-1)$ Freiheitsgraden, wobei $\Lambda' = \prod_{k=p+1}^q \frac{1}{1 + \lambda_k}$ ist und N, q die Anzahl der Probanden bzw. die Anzahl der Populationen darstellen. Der Anteil der gesamten Diskriminanzleistung, der in der k -ten Diskriminanzfunktion enthalten ist, kann dann als

$$(6.10) \quad \frac{\lambda_k}{\sum_{k=1}^q \lambda_k}$$

angenommen werden ([13], Kap.9). Aufgrund dieses Kriteriums

können solche Diskriminanzvariablen (-funktionen) y_k vernachlässigt werden, die eine geringe Diskriminanzleistung aufweisen.

6.5. Sequentielle Kovarianzanalyse

Man geht hier prinzipiell davon aus, den gesamten Variablensatz in zwei Untersätze zu verteilen: in einen Untersatz bestehend aus m_1 zur Populationstrennung zugelassenen Variablen und in einen Untersatz bestehend aus den restlichen $m_2 = m - m_1$ Variablen. Eine Kovarianzanalyse der beiden Untersätze wird durchgeführt. Dadurch wird der Einfluß des ersten Variablensatzes (Variablen, die eine wichtige Trenninformation enthalten) vom zweiten Variablensatz (zusätzliche Variablen, die auch redundant sein können) abgezogen. Die Residuen, d.h. der restliche Teil des letzten Variablensatzes wird anschließend analysiert.

Die Matrizen Σ und T sind entsprechend der Variablensätze durch folgende Partitionen darzustellen

$$(6.11) \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \quad T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix},$$

wobei Σ_{11}, T_{11} den Rang m_1 und Σ_{22}, T_{22} den Rang m_2 haben. Es folgt eine Berechnung der Residuen-Matrizen

$$(6.12) \quad \begin{aligned} \Sigma_{2 \cdot 1} &= \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \\ T_{2 \cdot 1} &= T_{22} - T_{21} T_{11}^{-1} T_{12} \end{aligned}$$

und des daraus resultierenden Wilks'-Lambda-Quotienten

$$(6.13) \quad \Lambda = \frac{|\Sigma_{2..}|}{|T_{2..}|}$$

Hier kann erneut Rao's F-näherung eingesetzt werden. Der Übergang einer Variablen vom zweiten Satz zum ersten wird dann nach dem Wert des univariaten F-Quotienten

$$(6.14) \quad F_j = \frac{(t_{2..j} - \sigma_{2..j})(q-1)}{\sigma_{2..j} (N - q - m_1)}$$

entschieden, wobei $t_{2..j}$ und $\sigma_{2..j}$ die j-ten Diagonalelemente der Matrizen $T_{2..}$ bzw. $\Sigma_{2..}$ sind. Dieser Quotient beantwortet die Frage ob die zusätzliche Variable j eine neue, nicht im alten Variablensatz enthaltene Information über die Verschiedenheit der Populationen bringt.

Die Methode der sequentiellen Kovarianzanalyse stützt sich auf den abhängigen F-Tests der Form (6.14). Die Variablen werden nacheinander in die Berechnung einbezogen, wobei die Zusammenhänge (6.12), (6.13) und (6.14) erneut berechnet werden müssen ([17], S.214a).

7. Informationstheoretische Diskriminanzanalyse

Im vorangegangenen Kapitel sind einige der wichtigsten Methoden zur Reduktion der Dimensionen in Klassifikationsaufgaben beschrieben worden. Das Prinzip der Hauptkomponenten und das Testen der Nullhypothese sind die zwei grundlegenden Prinzipien dieser Methoden. Es wird im weiteren versucht, die Reduktion der Dimensionen vom informationstheoretischen Gesichtspunkt aus zu betrachten. Das Konzept der "Trenninformation" zwischen zwei Populationen ist dafür erforderlich.

7.1. Das Konzept der "Trenninformation" und "informationstheoretischen Divergenz"

Der Begriff "Information" hat leider bei verschiedenen Autoren keine übereinstimmende Definition. Ein Versuch, Shannon's und Wiener's Definition zu verallgemeinern, wurde von Kullback und Leibler [34] unternommen. In derselben Arbeit zeigen beide Autoren, daß diese Definition im Zusammenhang mit Fisher's Informationsmaß steht.

Nach Kullback und Leibler ist die Trenninformation zwischen zwei Populationen als Logarithmus des Quotienten der Wahrscheinlichkeitsdichten definiert. Haben zwei Populationen D^1 und D^2 die Wahrscheinlichkeitsdichten $f_1(\mathbf{X})$ bzw. $f_2(\mathbf{X})$ vorausgesetzt, so schreibt sich die Trenninformation an der Stelle $\mathbf{x}$ unter Berücksichtigung von D^1 :

$$(7.1) \quad I_1(\mathbf{x}) = \log [f_1(\mathbf{x})/f_2(\mathbf{x})] \quad .$$

Die Summe der durchschnittlichen Trenninformationen unter

Berücksichtigung von Population D^1 und von Population D^2 ergibt die informationstheoretische Divergenz

$$(7.2) \quad J(1,2) = \int_S f_1(\mathbf{x}) \log \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} d\mathbf{x} + \int_S f_2(\mathbf{x}) \log \frac{f_2(\mathbf{x})}{f_1(\mathbf{x})} d\mathbf{x} ,$$

wobei die Integration über den gesamten Stichprobenraum stattfindet. Werden für D^1 und D^2 zwei normalverteilte Populationen $N(\boldsymbol{\mu}^1, \boldsymbol{\Sigma}^1)$ bzw. $N(\boldsymbol{\mu}^2, \boldsymbol{\Sigma}^2)$ vorausgesetzt, so erhält man für die Divergenz (7.2)

$$(7.3) \quad J(1,2) = \frac{1}{2} \text{tr} \{ (\boldsymbol{\Sigma}^1 - \boldsymbol{\Sigma}^2) [(\boldsymbol{\Sigma}^2)^{-1} - (\boldsymbol{\Sigma}^1)^{-1}] \} + \\ + \frac{1}{2} \boldsymbol{\delta}^T [(\boldsymbol{\Sigma}^1)^{-1} + (\boldsymbol{\Sigma}^2)^{-1}] \boldsymbol{\delta} ,$$

wobei $\boldsymbol{\delta} = \boldsymbol{\mu}^1 - \boldsymbol{\mu}^2$ ist und tr die Spur der Matrix bezeichnet. Für den Fall, daß $\boldsymbol{\Sigma}^1 = \boldsymbol{\Sigma}^2 = \boldsymbol{\Sigma}$ ist, folgt

$$(7.4) \quad J(1,2) = \boldsymbol{\delta}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\delta} ,$$

d.h. die informationstheoretische Divergenz ist gleich dem Mahalanobis-Abstand. Wenn die Varianz-Kovarianz-Matrix in den beiden Populationen unterschiedlich ist, kann die Divergenz (7.3) als Verallgemeinerung des Mahalanobis-Abstandes betrachtet werden.

Theorem (7.1): Vorgegeben sind zwei Stichproben $\mathbf{z}_k = (\mathbf{x}_k^i, (i=1,2,\dots,N))$, $k=1,2$ aus den normalverteilten Populationen $N(\boldsymbol{\mu}^1, \boldsymbol{\Sigma}^1)$ bzw. $N(\boldsymbol{\mu}^2, \boldsymbol{\Sigma}^2)$. Der aufgrund dieser Stichproben berechnete Wert

$$\hat{J}(1,2) = \frac{1}{2} \text{tr} \{ (\mathbf{S}^1 - \mathbf{S}^2) [(\mathbf{S}^2)^{-1} - (\mathbf{S}^1)^{-1}] \} + \\ + \frac{1}{2} \hat{\boldsymbol{\delta}}^T [(\mathbf{S}^1)^{-1} + (\mathbf{S}^2)^{-1}] \hat{\boldsymbol{\delta}} ,$$

wobei $\hat{\delta} = \bar{x}^1 - \bar{x}^2$ ist und die Matrizen S^1 und S^2 aus dem Zusammenhang (6.1) folgen, bestimmt dann den Maximum-Likelihood-Schätzer der informationstheoretischen Divergenz $J(1,2)$.

Beweis : Aus dem Zusammenhang (7.3) geht hervor, daß $J(1,2)$ eine eindeutige Funktion der Parameter $\delta = \mu^1 - \mu^2$, Σ^1 und Σ^2 ist. Entsprechend des Korollars (3.2.1.) vom [2] ist der Maximum-Likelihood-Schätzer für $J(1,2)$ durch dieselbe Funktion der Schätzungen $\hat{\delta} = \bar{x}^1 - \bar{x}^2$, $[(N-1)/N] S^1$ und $[(N-1)/N] S^2$ zu bestimmen. In diesem Fall stellt $\hat{J}(1,2)$ unter Annahme asymptotischer Äquivalenz ($N \rightarrow \infty$) den Maximum-Likelihood-Schätzer der informationstheoretischen Divergenz $J(1,2)$ dar. Q.E.D.

Aufgrund des oben ausgedrückten Theorems wird in dem experimentellen Teil dieser Arbeit der Maximum-Likelihood-Schätzer der informationstheoretischen Divergenz berechnet.

7.2. Informationstheoretische Erklärung der Hauptkomponentenmethode

Bei der Hauptachsentransformation wird unter der Voraussetzung $\delta = 0$ nach einer linearen Transformation

$$(7.5) \quad y = \alpha^T x$$

zum Maximieren der Divergenz (7.3) gesucht [33], wobei $\alpha^T = [\alpha_1, \alpha_2, \dots, \alpha_m]$ ein Koeffizientenvektor ist. Die Lö-

sung folgt aus der Matrizengleichung

$$\Sigma^1 = \lambda \Sigma^2, \quad ,$$

wobei λ einen Eigenwert der Determinantengleichung

$$(7.6) \quad |\Sigma^1 - \lambda \Sigma^2| = 0$$

darstellt. Unter der Annahme, daß die Gleichung (7.6) ungleiche, reelle und positive Lösungen λ_i hat, kann für jeden solchen Eigenwert λ_i ein Eigenvektor α_i berechnet werden.

Führt man die Transformation (7.5) mit Hilfe des Vektors α_i durch, dann ändert sich die Divergenz (7.3) folgendermaßen:

$$(7.7) \quad J'(1,2) = \frac{1}{2} \frac{\alpha_i^T \Sigma^1 \alpha_i}{\alpha_i^T \Sigma^2 \alpha_i} + \frac{1}{2} \frac{\alpha_i^T \Sigma^2 \alpha_i}{\alpha_i^T \Sigma^1 \alpha_i} - 1.$$

Nach der Vereinfachung gilt:

$$(7.8) \quad J'(1,2,\lambda_i) = \frac{1}{2} \lambda_i + \frac{1}{2\lambda_i} - 1.$$

Wenn $\{\alpha_i; (i=1,2,\dots,m)\}$ als Spalten einer $m \times m$ Matrix **A** berücksichtigt werden, können die Transformationen (7.5), für verschiedene i , im folgenden Zusammenhang zusammengefaßt werden

$$(7.9) \quad y = Ax.$$

Die Gesamtdivergenz der neuen Variablen y_i läßt sich durch Summation der einzelnen Variablendivergenzen zusammenschreiben

$$(7.10) \quad J'(1,2) = \sum_{i=1}^m J(1,2,\lambda_i).$$

Kullback [33] zeigt, daß die Hauptachsentransformation (7.9) die Gesamtdivergenz nicht ändert, d.h.

$$(7.11) \quad J(1,2) = J'(1,2).$$

Daraus schließt man, daß $J(1,2)$ eine Invariante für die Hauptachsentransformation ist und die Transformation zu unabhängigen (nicht korrelierten) Variablen die Bildung der Gesamtdivergenz aus einzelnen Summengliedern (Variablen-divergenzen) erlaubt.

Folgt nach der Hauptachsentransformation eine Reduktion der Dimensionen, so muß genau festgestellt werden, wieviel Trenninformation dabei verlorenggeht. Aufgrund dieser Tatsache werde ich im folgenden eine kritische Diskussion der Hauptkomponentenmethode mit dem Ziel durchführen, eine wichtige Einschränkung dieser Methode aufzuzeigen.

Aus dem oben beschriebenen Formalismus ist zu erkennen, daß bei der Berechnung der Hauptkomponenten der Vektor δ gleich Null gesetzt wird. Mit anderen Worten, die zwei Populationen werden von Anfang an als konzentrisch betrachtet. Es können auch solche Variablen vernachlässigt werden, die eine geringe Varianz haben aber einen wichtigen Beitrag zum Zentrenabstand der Populationen bringen. Dabei kann ein erheblicher Verlust an Trenninformation auftreten. In den Trennalgorithmen macht sich dann eine wesentliche Steigerung der Fehlerquote bemerkbar.

Um diese bereits geschilderte Einschränkung der Hauptkomponentenmethode näher zu untersuchen, werden Monte-Carlo-Experimente geplant. Bei allen Experimenten werden zwei fast konzentrische und eine dritte stark exzentrische Population beobachtet (§8.). Man erhält die Exzentrizität der dritten Population durch Verschiebung des Mittelwertes bei einer Variablen mit geringer Varianz. Nach der Durchführung der Hauptachsentransformation folgt eine Reduktion der Dimensionen. Damit ist es möglich, das Versagen der Hauptkomponentenmethode bei exzentrischen Populationen zu beweisen. Als Gütekriterium der Dimensionsreduktion wird die Fehlklassifikationsrate der mit den bereits reduzierten Daten durchgeführten Re- und Neuklassifikationen berücksichtigt.

7.3. Die informationstheoretische Divergenz als Kriterium für die Dimensionenreduktion

Ein Hauptziel der vorliegenden Arbeit ist, ein informationstheoretisches Maß - die Variablendivergenz - als Kriterium für die Reduktion der Dimensionen anzugeben. Dazu ist es notwendig, zu beweisen, daß für eine Transformation der Form (7.9) der Zusammenhang (7.10) auch für exzentrische Populationen ($\delta \neq 0$) gültig bleibt.

Der Zwei-Populationen-Fall. Falls $\delta \neq 0$ ändert sich die Divergenz (7.3) nach der Transformation mit dem Eigenvektor α_1 , wie folgt

$$(7.12) \quad J^*(1,2,\lambda_1) = \frac{1}{2} \frac{\alpha_1^T \Sigma^1 \alpha_1}{\alpha_1^T \Sigma^2 \alpha_1} + \frac{\alpha_1^T \Sigma^2 \alpha_1}{\alpha_1^T \Sigma^1 \alpha_1} - 1 + \\ + \frac{1}{2} (\alpha_1^T \delta)^T \left[\frac{1}{\alpha_1^T \Sigma^1 \alpha_1} + \frac{1}{\alpha_1^T \Sigma^2 \alpha_1} \right] (\alpha_1^T \delta) .$$

Unter der Annahme, daß $\lambda_1 > 0$ und reell ist und für den entsprechenden Eigenvektor α_1 die Bedingung $\alpha_1^T \Sigma^2 \alpha_1 = 1$ gilt, erhält man für (7.12)

$$(7.13) \quad J^*(1,2,\lambda_1) = \frac{1}{2} \lambda_1 + \frac{1}{2\lambda_1} - 1 + \frac{1}{2} (\alpha_1^T \delta)^2 \left[\frac{1}{\lambda_1} + 1 \right] .$$

Die Gleichung (7.13)

drückt die Variablendivergenz der neuen Variablen y_1 aus. Dieses Ergebnis ist zu erwarten, da eine Hauptachsentransformation die Populationen in einen Raum mit unabhängigen (nicht korrelierten) Variablen überträgt, und zwar sowohl für $\delta = 0$, als auch für $\delta \neq 0$.

Eine weitere Anforderung die man hier stellen muß, ist der Beweis, daß die Gesamtdivergenz aus einzelnen Summengliedern (Variablendivergenzen) gebildet und Kullback's

Gleichung (7.10) für den Fall $\delta \neq 0$ erhalten werden kann. Hierzu ist dann zu prüfen, ob die Summe des letzten Terms der Gleichung (7.12), für i , den letzten Term der Gleichung (7.3) ergibt. Wenn A die orthogonale Matrix ist, deren Spalten aus den Eigenvektoren α_i , ($i = 1, 2, \dots, m$) bestehen, dann läßt sich diese Summe durch

$$\sum_i (\alpha_i^T \delta)^T \left[\frac{1}{\alpha_i^T \Sigma^1 \alpha_i} + \frac{1}{\alpha_i^T \Sigma^2 \alpha_i} \right] (\alpha_i^T \delta) = \delta^T A [A^T (\Sigma^1)^{-1} A + A^T (\Sigma^2)^{-1} A] A^T \delta = \delta^T [(\Sigma^1)^{-1} + (\Sigma^2)^{-1}] \delta$$

beschreiben. Damit ist bewiesen, daß

$$(7.14) \quad J(1,2) = \sum_i J^*(1,2,\lambda_i)$$

gilt, wobei $J^*(1,2,\lambda_i)$ die im Falle $\delta \neq 0$ nach der Hauptachsentransformation erhaltene Divergenz der Variablen y_i darstellt.

Theorem (7.2): Vorgegeben sind zwei Populationen D^1 und D^2 mit $\Sigma^1 = \Sigma^2 = \Sigma$ und $\delta \neq 0$. Das Kriterium der minimalen Variablendivergenz ist dann für eine Reduktion der Dimensionen optimal, da es die Fehlklassifikationsrate minimiert.

Beweis : Für den Fall, daß $\Sigma^1 = \Sigma^2 = \Sigma$ und $\delta \neq 0$ ist, schreibt sich der Mahalanobis-Abstand zwischen den Populationen D^1 und D^2 folgendermaßen

$$d^2 = J(1,2) = \delta^T \Sigma^{-1} \delta$$

Die Fehlklassifikationsrate ist - wie schon im Paragraphen (5.1) beschrieben wurde - durch die Funktion G der positiven Wurzel des Mahalanobis-Abstandes ausgedrückt:

$$\alpha_1(\xi) = G\left(-\frac{1}{2}d\right)$$

Eine genaue Beobachtung der Funktion G zeigt, daß die Fehlklassifikationsrate sinkt, wenn d steigt. Nach der Hauptachsentransformation (7.9) erhält man für den Mahalanobis-Abstand

$$d^2 = J^*(1,2) = \boldsymbol{\delta}^T \mathbf{A} \mathbf{L}^{-1} \mathbf{A}^T \boldsymbol{\delta} \quad ,$$

wobei $\mathbf{L} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_m]$ ist und λ_i ($i=1,2,\dots,m$) Eigenwerte der Matrix $\boldsymbol{\Sigma}$ darstellen. Ausführlicher schreibt man

$$(7.15) \quad d^2 = J^*(1,2) = \sum_{i=1}^m \frac{(\boldsymbol{\alpha}_i^T \boldsymbol{\delta})^2}{\lambda_i} \quad .$$

Wird jetzt die Variable 1 , $1 \in \{1,2,\dots,m\}$ reduziert, so verringert sich der Abstand d^2 mit dem 1 -ten Term der Summe (7.15). Um eine optimale Reduktion der Dimensionen durchzuführen (eine minimale Fehlklassifikationsrate zu erreichen) muß für 1 folgende Bedingung erfüllt werden

$$J^*(1,2,\lambda_e) = \frac{(\boldsymbol{\alpha}_e^T \boldsymbol{\delta})^2}{\lambda_e} = \min \left\{ \frac{(\boldsymbol{\alpha}_i^T \boldsymbol{\delta})^2}{\lambda_i} , (i=1,2,\dots,m) \right\} ,$$

d.h. die reduzierte Variable muß eine minimale Divergenz haben. Q.E.D.

Folgt die Reduktion der Variablen 1 aufgrund der minimalen Varianz

$$\lambda_e = \min \{ \lambda_1, \lambda_2, \dots, \lambda_m \} \quad ,$$

so wird der Term

$$\frac{(\boldsymbol{\alpha}_e^T \boldsymbol{\delta})^2}{\lambda_e} = \frac{(\boldsymbol{\alpha}_e^T \boldsymbol{\delta})^2}{\boldsymbol{\alpha}_e^T \boldsymbol{\Sigma} \boldsymbol{\alpha}_e}$$

kaum minimiert. Man geht hier davon aus, daß der Quotient aus der quadrierten linearen Form $(\boldsymbol{\alpha}_e^T \boldsymbol{\delta})^2$ (die lineare Form stellt für $\boldsymbol{\delta} = \text{konstant}$ eine Hyperebene dar, die

Koordinatenursprung enthält) und der entsprechenden quadratischen Form $\lambda_e = \alpha_e^T \Sigma \alpha_e$ für das Minimum der quadratischen Form in der Regel nicht minimiert wird.

Damit wurde bewiesen, daß die Variablendivergenz ein optimales Kriterium für die Reduktion der Dimensionen bildet, falls die Varianz-Kovarianz-Matrix gleich in den beiden Populationen ist. Ist diese Matrix in den beiden Populationen dagegen unterschiedlich, so kann die übliche Form des Mahalanobis-Abstandes nicht verwendet werden. Die Ableitung der Fehlklassifikationsrate $\alpha_e(\xi)$ folgt üblicherweise aufgrund der Hypothese, daß $\Sigma^1 = \Sigma^2 = \Sigma$ ist und eine Verteilung ist bis jetzt für den allgemeinen Fall unbekannt.

Aufgrund der in ([33], S.90) beschriebenen Eigenschaften der Divergenz bildet der Quotient

$$(7.16) \quad e_i = \frac{J^*(1,2,\lambda_i)}{J(1,2)}$$

auch für den Fall $\delta \neq 0$ ein Effizienzmaß der Variablen y_i . Dieses Maß ist dann die in der i-ten Variablen enthaltenen Trenninformation für die Trennung zwischen den Populationen D^1 und D^2 .

Aufgrund des Theorems (7.2) ist zu schließen, daß die Variablendivergenz das beste Auswahlkriterium bei der Reduktion der Dimensionen bildet, wenn es um die Minimierung der Fehlklassifikationsrate geht (Für den Fall $\Sigma^1 = \Sigma^2 = \Sigma$). Um dasselbe für ungleiche Varianz-Kovarianz-Matrizen in den beiden Populationen zu beweisen, werden im folgenden Monte-Carlo-Experimente konstruiert. Das oben beschriebene Kriterium für die Dimensionenreduktion

wurde bei den in dieser Arbeit beschriebenen Monte-Carlo-Experimenten zugrunde gelegt und wird im weiteren als "Prinzip der minimalen Variablendivergenz" bezeichnet.

Theorem (7.3): Vorgegeben sind zwei Stichproben $\mathbf{x}_k = \{\mathbf{x}_k^i, (i=1,2,\dots,N)\}$, $k=1,2$ aus den normalverteilten Populationen $N(\boldsymbol{\mu}^1, \boldsymbol{\Sigma}^1)$ bzw. $N(\boldsymbol{\mu}^2, \boldsymbol{\Sigma}^2)$. Der aufgrund dieser Stichproben berechnete Wert

$$(7.17) \quad \hat{J}(1,2,\hat{\lambda}_1) = \frac{1}{2}\hat{\lambda}_1 + \frac{1}{2\hat{\lambda}_1} - 1 + \frac{1}{2}(\hat{\boldsymbol{\alpha}}_1^T \hat{\boldsymbol{\delta}})^2 \left[\frac{1}{\hat{\lambda}_1} + 1 \right],$$

wobei $\hat{\lambda}_1$ und $\hat{\boldsymbol{\alpha}}_1$ ($i=1,2,\dots,m$) Lösungen des Gleichungssystems

$$(7.18) \quad \begin{aligned} |\mathbf{S}^1 - \lambda \mathbf{S}^2| &= 0 \\ (\mathbf{S}^1 - \hat{\lambda}_1 \mathbf{S}^2) \hat{\boldsymbol{\alpha}}_1 &= \mathbf{0} \quad (i=1,2,\dots,m) \\ \hat{\boldsymbol{\alpha}}_1^T \mathbf{S}^2 \hat{\boldsymbol{\alpha}}_1 &= 1 \end{aligned}$$

darstellen, bestimmt dann den Maximum-Likelihood-Schätzer der Variablendivergenz $J^*(1,2,\lambda_1)$.

Beweis: Entsprechend des Theorems (11.3.1.) vom [2] sind die Maximum-Likelihood-Schätzungen $\hat{\lambda}_1$, $\hat{\boldsymbol{\alpha}}_1$ ($i=1,2,\dots,m$) durch dasselbe Gleichungssystem (7.18) von den Schätzungen $\bar{\mathbf{x}}^1$, $\bar{\mathbf{x}}^2$, $[(N-1)/N] \mathbf{S}^1$ und $[(N-1)/N] \mathbf{S}^2$ zu erhalten. Die mit Hilfe der Schätzungen $\hat{\lambda}_1$, $\hat{\boldsymbol{\alpha}}_1$ ($i=1,2,\dots,m$) und $\hat{\boldsymbol{\delta}} = \bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2$ aufgrund der eindeutigen Funktion (7.17) berechnete Variablendivergenz ist dann dem Maximum-Likelihood-Schätzer der Variablendivergenz $J^*(1,2,\lambda_1)$ asymptotisch äquivalent (s.[2], Korollar (3.2.1.)). Q.E.D.

Die Größe e_1 liefert ein genaues Maß der bei der Reduktion der Dimensionen verlorengegangenen Trenninformation.

Das Maß

$$e_i = \sum_{j=1}^{m_i} e_{ij} \quad , \quad m_i < m$$

bestimmt die in den m_i selektierten Variablen gebliebene Trenninformation. Seine Schätzung kann mit Hilfe des Maximum-Likelihood-Schätzer der Variablen-divergenz berechnet werden.

Mehr als zwei Populationen. In diesem Fall wird üblicherweise die Total-Matrix T aller Populationen nach den Hauptachsen transformiert. Man betrachtet die Gesamtheit aller Individuen als eine einzige Population mit dem Zentrum im Hauptzentrum der Populationen. Die Hauptachsen-transformation bringt dieses Mal keine Unabhängigkeit der Variablen voneinander in den einzelnen Populationen und die Gesamtdivergenz für zwei willkürlich ausgewählte Populationen kann nicht mehr aus den Summengliedern $J^*(1,2,\lambda_i)$ gebildet werden. Man kann jedoch den Zwei-Populationen-Formalismus für die Populationen

$$D^i \quad \text{und} \quad D = \bigcup_{j \neq i} D^j \quad ; \quad i=1,2,\dots,q$$

verwenden. Die Berechnung wird q -Mal für jede reduzierte Variable wiederholt. In jedem Reduktionsschritt wird D^i derart ausgewählt, daß die Trenninformation der reduzierten Variable minimal wird.

Mit Hilfe dieses Algorithmus können, aufgrund einer Reduktion der Dimensionen, die Populationen auseinandergelassen und die Fehlklassifikationsrate erheblich verbessert werden. Aufgrund von apriori Informationen, ist es möglich, die Reduktion der Dimensionen gezielt bezüglich einer oder zweier Populationen durchzuführen.

8. Untersuchungen mit Hilfe von Monte-Carlo-Experimenten

Die Monte-Carlo-Zufallsexperimente werden in zwei Etappen durchgeführt:

- (1) Aufbau des Datenmaterials mit Hilfe eines Datenzufalls-generators und
- (2) Durchführung der Klassifikationen mit den reduzierten Daten, wobei einmal zur Reduktion der Dimensionen das Prinzip der minimalen Variablenvarianz und zum anderen das Prinzip der minimalen Variablendivergenz verwendet wird.

In der ersten Etappe werden mit Hilfe eines Simulationssystems m -dimensionale normal-verteilte Zufallsvektoren erzeugt. Zu erstellen sind drei normalverteilte Populationen, wobei die bei der Erzeugung von mir angegebenen Modellparameter unter bestimmten Bedingungen (§ 8.1.) ausgewählt werden.

In der zweiten Etappe schätzt man zunächst die Parameter der Populationen, dann folgt die Reduktion der Dimensionen und schließlich wird die Klassifikation durchgeführt.

Die Experimente beginnen mit dem Erstellen einer Stichprobe von bereits diagnostizierten Patienten (erhärtete Fälle). Danach wird diese Stichprobe zum Abschätzen der Parameter und für Reklassifikationen verwendet.

Für die Neuklassifikationen ist es nötig, eine zusätzliche unabhängige Stichprobe mit undiagnostizierten Patienten zu erstellen. Die Patientendaten dieser Stichprobe werden nur zur Klassifikation verwendet, ohne daß sie auch zum Abschätzen der Parameter beitragen.

8.1. Aufbau des Simulationssystems und Beschreibung des Datenmaterials

Das Simulationssystem braucht bei der Erzeugung der Zufallsvektoren bestimmte Modellparameter, d.h. die Mittelpunktsvektoren und Varianz-Kovarianz-Matrizen der zu erstellen- den Populationen. Die Modellparameter werden so ausgewählt, daß die Reduktionsmethoden von drei Gesichtspunkten aus getestet werden können:

- (1) die Untersuchung des Verhaltens der Methode bei unterschiedlicher Anzahl der Variablen,
- (2) Feststellung des Einflusses der unterschiedlich großen Stichproben auf die Reduktion der Dimensionen, und
- (3) Abschätzung der Empfindlichkeit der Methode bei unterschiedlichem Abstand zwischen den Populationszentren.

Zum ersten Aspekt kann man von der Theorie her bereits erwarten, daß je größer die Anzahl der Dimensionen ist, desto wirksamer die Methode sein wird. Um Rechenzeit und -platz zu sparen, wird zunächst die Reduktion von 10 auf 5 Dimensionen durchgeführt. Die Reduktion von 60 auf 50 Dimensionen benötigt die 15-fache Rechenzeit. Bemerkenswert ist der Fall, bei dem die Anzahl der Probanden mit der Anzahl der Variablen vergleichbar ist. Wie im §5.2. beschrieben wird ¹⁾, kann in solchen Fällen eine Reduktion der Dimensionen mehr schaden als nützen. Infolgedessen, ist bei Experimenten mit 60 Variablen und relativ geringen Stichprobenmengen eine zusätzliche Überprüfung der gesamten Trenninformation notwendig.

Im Zusammenhang mit dem zweiten Aspekt, werden Experimente mit Stichproben von 150, 300 und 900 Probanden pro Population durchgeführt. Für jede Stichprobengröße wird die Fehlerquote der Klassifikationen als Mittelwert über

1) siehe Tabelle 1 ($n_1=n_2=50$ und $m=50$)

Tabelle 2 (Modellparameter für je 3 Populationen mit 10 Variablen)

		Mittelwerte μ_i^k									
Pop.k	Vari	1	2	3	4	5	6	7	8	9	10
1		30.0	5.0	7.0	2.0	9.0	3.0	10.0	10.0	3.0	0.5
2		4.0	6.0	8.3	3.0	7.0	6.0	11.0	10.0	5.5	6.0
3		2.0	4.0	6.3	1.0	7.0	8.0	10.0	12.0	0.5	4.0

		Varianzen σ_i^k									
Pop.k	Vari	1	2	3	4	5	6	7	8	9	10
1		1.3	29.0	11.5	15.0	11.0	29.8	25.2	10.5	18.9	9.7
2		5.2	28.7	9.1	12.5	19.0	29.0	8.9	3.0	9.7	30.8
3		10.8	29.5	28.5	11.0	2.5	10.0	29.2	11.2	30.2	2.9

Mahalanobis - Abstände

$$d_{12}^2 = 14.1 \quad d_{13}^2 = 16.35 \quad d_{23}^2 = 0.25$$

Tabelle 3 (Modellparameter für je 3 Populationen mit 60 Variablen)

Mahalanobis - Abstände: $d_{12}^2 = 11.1$, $d_{13}^2 = 9.1$, $d_{23}^2 = 8.2$

μ^1	1	3.0	5.0	7.0	2.0	9.0	3.0	10.0	10.0	3.0	0.5	3.0	10.0	3.0	5.0	7.0	2.0	10.0	9.0	3.0	0.5
	2	10.0	3.0	3.0	5.0	2.0	7.0	10.0	3.0	9.0	0.5	0.5	10.0	9.0	2.0	3.0	3.0	3.0	7.0	5.0	10.0
	3	10.0	0.5	3.0	10.0	3.0	3.0	2.0	9.0	5.0	7.0	12.5	3.0	6.5	2.0	3.4	-0.5	0.0	13.3	3.5	5.2
μ^2	1	4.0	6.0	8.3	3.0	7.0	6.0	11.0	10.0	5.5	6.0	6.0	10.0	4.0	6.0	8.3	3.0	11.0	7.0	5.5	6.0
	2	10.0	6.0	4.0	6.0	3.0	8.3	11.0	5.5	7.0	6.0	6.0	10.0	7.0	3.0	6.0	4.0	5.5	8.3	6.0	11.0
	3	11.0	6.0	5.5	10.0	4.0	6.0	3.0	7.0	6.0	8.3	10.0	3.0	5.0	3.0	1.5	3.8	2.0	7.9	4.0	5.0
μ^3	1	2.0	4.0	6.3	1.0	7.0	8.0	10.0	12.0	0.5	4.0	8.0	12.0	2.0	4.0	6.3	1.0	10.0	7.0	0.5	4.0
	2	12.0	8.0	2.0	4.0	1.0	6.3	10.0	0.5	7.0	4.0	4.0	12.0	7.0	1.0	8.0	2.0	0.5	6.3	4.0	10.0
	3	10.0	4.0	0.5	12.0	2.0	8.0	1.0	7.0	4.0	6.3	15.0	2.3	8.7	2.8	5.2	6.1	0.5	11.0	3.0	5.5
λ^1	1	1.3	29.0	11.5	15.0	11.0	29.8	25.2	10.5	18.9	9.7	5.2	28.7	9.1	12.2	19.0	29.0	8.9	3.0	9.7	30.8
	2	10.8	29.5	28.5	11.0	2.5	10.0	29.2	11.2	30.2	2.9	3.0	5.0	7.0	2.0	9.0	3.0	10.0	10.0	3.0	0.5
	3	3.0	5.0	7.0	2.0	9.0	3.0	10.0	10.0	3.0	0.5	6.0	10.0	7.0	3.0	6.0	4.0	5.5	8.3	6.0	11.0
λ^2	1	10.0	6.0	4.0	6.0	3.0	8.3	11.0	5.5	7.0	6.0	10.0	3.0	3.0	5.0	2.0	7.0	10.0	3.0	9.0	0.5
	2	4.0	12.0	7.0	1.0	8.0	2.0	0.5	6.3	4.0	10.0	0.5	10.0	9.0	2.0	3.0	3.0	3.0	7.0	5.0	10.0
	3	10.0	4.0	0.5	12.0	2.0	8.0	1.0	7.0	4.0	6.3	11.0	6.0	5.5	10.0	4.0	6.0	3.0	7.0	6.0	8.3
λ^3	1	10.0	3.0	5.0	3.0	1.5	3.8	2.0	7.9	4.0	5.0	8.7	5.2	2.8	6.1	2.3	15.0	0.5	5.5	11.0	3.0
	2	5.0	1.5	3.0	3.8	3.0	10.0	2.0	5.0	7.9	4.0	4.0	1.5	5.0	3.0	3.0	3.8	2.0	5.0	10.0	7.9
	3	11.0	3.0	8.7	2.8	0.5	6.1	2.3	15.0	11.0	5.5	7.9	4.0	5.0	3.0	2.0	3.8	3.0	10.0	7.9	5.0

Tabelle 4 (Transformationsmatrix)

	1	2	3	4	5	6	7	8	9	10
1	.20	.07	.15	.21	.25	.10	.28	.22	.15	.08
2		.15	.20	.18	.11	.08	.16	.19	.10	.23
3			.13	.15	.30	.20	2.08	.08	.22	.27
4				.22	.25	.18	.23	.12	.25	.17
5					.18	.20	.19	.30	.26	.27
6						.15	.22	.18	.18	.31
7							.22	.19	.21	.08
8								.19	.27	.71
9									.21	.02
10										.28

Tabelle 5 Fehlklassifikationsrate in Prozente

Mittelwerte und Standardabweichungen für

150 Probanden pro Stichprobe				300 Probanden pro Stichprobe				900 Probanden pro Stichprobe			
REKLASSIFIKATION	OHNE REDUKTION	REDUKTION AUFGRUND DER VARIANZ	REDUKTION AUFGRUND DER DIVERGENZ	OHNE REDUKTION	REDUKTION AUFGRUND DER VARIANZ	REDUKTION AUFGRUND DER DIVERGENZ	OHNE REDUKTION	REDUKTION AUFGRUND DER VARIANZ	REDUKTION AUFGRUND DER DIVERGENZ		
	P1	0	13 ± 3	0	14 ± 6	0	0	15 ± 4	0		
	P2	38 ± 6	49 ± 5	42 ± 5	38 ± 4	51 ± 5	41 ± 5	37 ± 2	58 ± 4	41 ± 3	
	P3	42 ± 6	52 ± 4	44 ± 4	43 ± 3	47 ± 4	43 ± 3	43 ± 1	45 ± 1	44 ± 1	
	Mittelwert	26,67	38,00	28,67	27,00	37,33	28,00	26,67	39,33	28,33	
NEUKLASSIFIKATION	P1	0	15 ± 4	0	15 ± 6	0	0	14 ± 2	0		
	P2	42 ± 7	52 ± 7	46 ± 7	42 ± 5	52 ± 4	46 ± 5	39 ± 4	57 ± 6	42 ± 5	
	P3	43 ± 4	50 ± 9	44 ± 7	46 ± 4	47 ± 5	46 ± 6	43 ± 3	45 ± 4	43 ± 3	
	Mittelwert	28,33	39,00	30,00	29,33	38,00	30,67	27,33	38,67	28,33	

Tabelle 6 Fehlklassifikationsrate in Prozente

Mittelwerte und Standardabweichungen für 150
Probanden pro Stichprobe (Datenmaterial mit
Interaktionen)

POPULATION		OHNE REDUKTION	REDUKTION AUFGRUND DER VARIANZ	REDUKTION AUFGRUND DER DIVERGENZ
REKLASSIFIKATION	P1	0	33 ± 4	0
	P2	38 ± 7	55 ± 5	42 ± 5
	P3	43 ± 5	47 ± 5	44 ± 4
	Mittelwert	27,00	45,00	28,67
NEUKLASSIFIKATION	P1	0	37 ± 7	0
	P2	43 ± 6	57 ± 8	46 ± 6
	P3	43 ± 4	44 ± 7	44 ± 6
	Mittelwert	28,67	46,00	30,00

10 Zufallsexperimenten mit unabhängigen Stichproben. Außer der Stichprobe von z.B. 150 Probanden, die die erhärteten Fälle der jeweils 3 Populationen bilden, wird bei jedem Experiment eine gleich große Stichprobe (in diesem Fall 150) mit einer Mischung von neuen Patienten aus allen Populationen für Neuklassifikationen erstellt. Bei den Experimenten mit Stichproben von 150 Probanden können auch zweidimensionale Projektionen des Diskriminanzraumes mit Hilfe eines Druckprogramms gedruckt werden.

Um die Empfindlichkeit der Reduktionsmethoden bei unterschiedlichem Abstand zwischen den Populationszentren zu untersuchen, werden bei allen Experimenten zwei fast konzentrische Populationen und eine dritte stark exzentrische Population beobachtet. Damit ist es möglich, die nur in der Varianz-Kovarianz-Matrix enthaltene Trenninformation bei konzentrischen Populationen und die gesamte Trenninformation bei exzentrischen Populationen zu untersuchen. In der Tabelle 2 und 3 sind die Modellparameter der Populationen und die Mahalanobis-Abstände zwischen den Populationen für 10 bzw. 60 Variablen abgebildet. Wie aus den Tabellen zu ersehen ist, sind die Variablen als voneinander unabhängig (mittels der Varianz-Diagonal-Matrizen) angegeben.

Bei der Auswahl der Modellparameter wird auch die Entropie der Trenninformationsmaße (7.16) für je zwei Populationen D^p und D^q berücksichtigt:

$$H_{p,q} = - \sum_{i=1}^{10} \frac{J^*(p,q,\sigma_i)}{J(p,q)} \ln \frac{J^*(p,q,\sigma_i)}{J(p,q)} .$$

Es wird dabei geachtet, daß konzentrische Populationen und exzentrische Populationen eine hohe Entropie ($H_{2,3}=1,96$) bzw. eine geringe Entropie ($H_{1,2}=0,33$, $H_{1,3}=0,49$) der Trenninformationsmaße haben, wobei $0 \leq H_{p,q} \leq 2,3$ ist. Dadurch ist es

möglich die erforderlichen Experimente auf ein Minimum zu reduzieren.

Für die 10-dimensionalen Populationen wird auch eine vollständige Varianz-Kovarianz-Matrix verwendet (s. Tabelle 4). Damit werden künstlich Interaktionen zwischen den einzelnen Variablen erzeugt. Hier ist zu bemerken, daß die Varianz-Kovarianz-Matrix in den durchgeführten Experimenten nicht mit Hilfe von Zufallszahlen aufgebaut wird. Dadurch werden solche Schwierigkeiten umgangen, die beim Aufbau von allgemeinen Kovarianzmatrizen durch Vorgabe von Korrelationskoeffizienten zwischen je zwei Variablen auftreten können ([12], S. 21). Man geht hier davon aus, daß die Interaktionen (Abhängigkeiten) von Anfang an fixiert sind und sich während der Experimente nicht ändern.

8.2. Darstellung und Diskussion der Ergebnisse

8.2.1. Überprüfung der Hauptkomponentenmethode

Diese Reihe von Experimenten werden mit dem Ziel durchgeführt, die Hauptkomponentenmethode vom informationstheoretischen Gesichtspunkt zu überprüfen. Entsprechend dieser Methode, werden solche Variablen als unwichtig für den Trennprozeß betrachtet, die in den Zufallsexperimenten ungefähr als Konstanten auftreten, d.h. sie haben eine geringe Varianz.

Im §7.2. wird beschrieben, daß die Hauptachsentransformation einen wesentlichen Nachteil hat: sie setzt voraus, daß die Populationen im Stichprobenraum konzentrisch sind. Dadurch können im Verlauf der Dimensionsreduktion zwangsläufig auch solche Variablen vernachlässigt werden, die zwar eine geringe Varianz haben, aber eine wichtige Information über den Zentrenabstand der Populationen mit sich bringen. Die Vernachlässigung dieser Variablen kann selbstverständlich einen erheblichen Verlust an Trenninformation zur Folge haben.

Um dieses zu überprüfen, werden zuerst durch Monte-Carlo-Experimente drei normalverteilte Populationen erstellt. Man wählt die Mittelwerte so aus, daß die drei Populationen am Anfang fast konzentrisch sind. Danach wird aber der Mittelwert einer Variablen mit einer relativ geringen Varianz in der ersten Population stark verschoben.

Eine multivariate Diskriminanzanalyse des wie oben aufgebauten Stichprobenraumes wird mit Hilfe des im DRZ Darmstadt entwickelten Programm DISK durchgeführt. Nach der

Diskriminanzanalyse folgt die Projektion der Daten in den Diskriminanzraum.

Die Abbildung 6 zeigt zwei-dimensionale Projektionen des Diskriminanzraumes. In dem oberen Teil dieser Abbildung, in dem der unreduzierte Diskriminanzraum dargestellt ist, kann eine ausgeprägte Verschiebung der ersten Population beobachtet werden. Nach einer Reduktion von 10 auf 5 Dimensionen mit Hilfe der Hauptachsentransformation entsteht der Raum, dessen Projektion in dem unteren Teil der Abbildung dargestellt ist. Dabei ist die Variable mit dem veränderten Mittelwert verlorengegangen.

In der Tabelle 5 sind die Fehlklassifikationsraten eingetragen. Bemerkenswert ist, daß bei Neuklassifikationen durchschnittlich 28% Fehler ohne jene Reduktion der Dimensionen und 39% Fehler mit Hauptachsentransformation gefunden werden. Die hohe Fehlerquote ist durch die geringe Exzentrizität aller Populationen, insbesondere der Populationen 2 und 3 zu erklären. Diese Populationen werden fast konzentrisch erhalten, um den Einfluß des Zentrenabstandes auf die Fehlklassifikationsrate beobachten zu können. Die Fehlerquote von 39% entspricht also dem unteren Teil der Abbildung 6, wobei die wichtigste Trennvariable verlorengegangen ist. Es ist hier darauf zu achten, daß die Populationen von unabhängigen Variablen bestimmt werden.

Falls Abhängigkeiten zwischen den Variablen vorliegen, sind auch andere Variablen vom geänderten Mittelwert betroffen. Damit wird die, durch den geänderten Mittelwert erhaltene Trenninformation den stark abhängigen Variablen übertragen. Ein Teil dieser Variablen geht bei der Reduktion der Dimensionen verloren. Unter diesen Bedingungen sind bei Neuklassifikationen durchschnittlich 29% Fehler ohne Reduktion der Dimensionen und 46% Fehler

mit Hauptachsentransformation zu erwarten (s.Tab.6,Abb.7).

Durch das Experiment wurde bestätigt, daß die Reduktion der Dimensionen aufgrund der Variablenvarianz ungeeignet ist, und zwar deswegen, weil dadurch ein Teil der Trenninformation verlorengeht.

8.2.2. Informationstheoretische Reduktion der Dimensionen

Aus dem Kapitel 7 geht hervor, daß vom informationstheoretischen Gesichtspunkt, die Variablendivergenz und nicht die Variablenvarianz als Auswahlkriterium der Variablen bei der Reduktion der Dimensionen zugrundegelegt werden muß. Dieses Verfahren ermöglicht sowohl eine gleichzeitige einfache Berechnung aller Variablendivergenzen, als auch die Ableitung eines exakten Maßes der verlorengegangenen Trenninformation.

Bei der Reduktion der Dimensionen ist das Verfahren durch Permutationen für jede Population anzuwenden, wobei die restlichen zwei Populationen als eine einzige einheitliche Population betrachtet wird.

Die Ergebnisse der durchgeführten Monte-Carlo-Experimente unter Anwendung der informationstheoretischen Methode zur Reduktion der Dimensionen können am besten anhand der Abbildung 8 interpretiert werden. Erzeugt wurden mit denselben Daten, wie in den oben beschriebenen Experimenten, zwei-dimensionale Projektionen des Diskriminanzraumes nach der Reduktion. Im oberen Teil der Abbildung 8 ist wie in der Abbildung 6 derselbe, mit Hilfe der Hauptachsentransformation erzeugte fünf-dimensionale Diskriminanzraum zu sehen. Zum Vergleich ist unten eine zwei-dimensionale Projektion des Diskriminanzraumes nach einer informations-

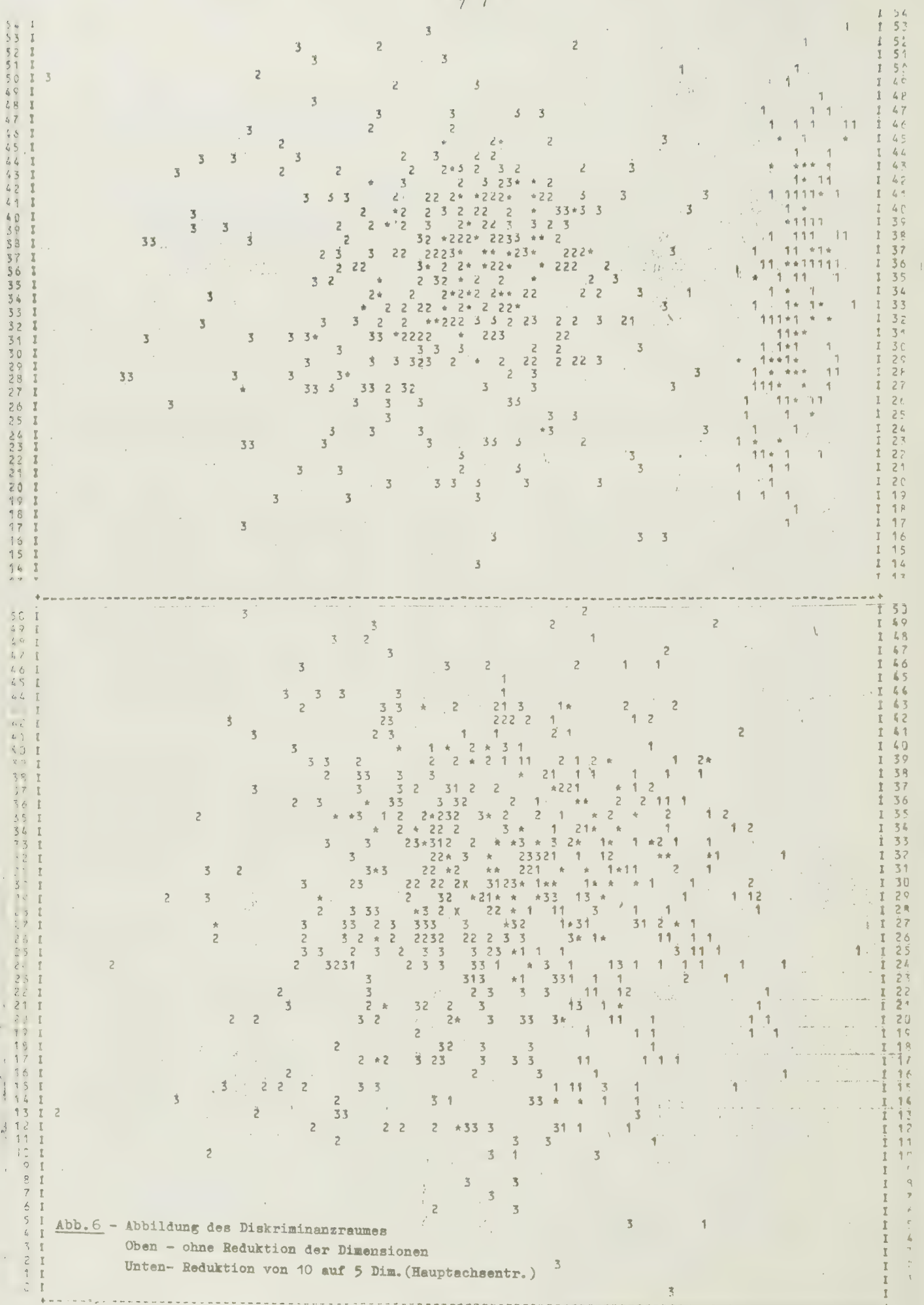


Abb.6 - Abbildung des Diskriminanzraumes
Oben - ohne Reduktion der Dimensionen
Unten- Reduktion von 10 auf 5 Dim.(Hauptachsentr.)

0123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890



Abb. 7 - Abbildung des Diskriminanzraumes (Abhängigkeiten)

Oben - ohne Reduktion der Dimensionen

Unten - Reduktion von 10 auf 5 Dim. (Hauptachsentr.)

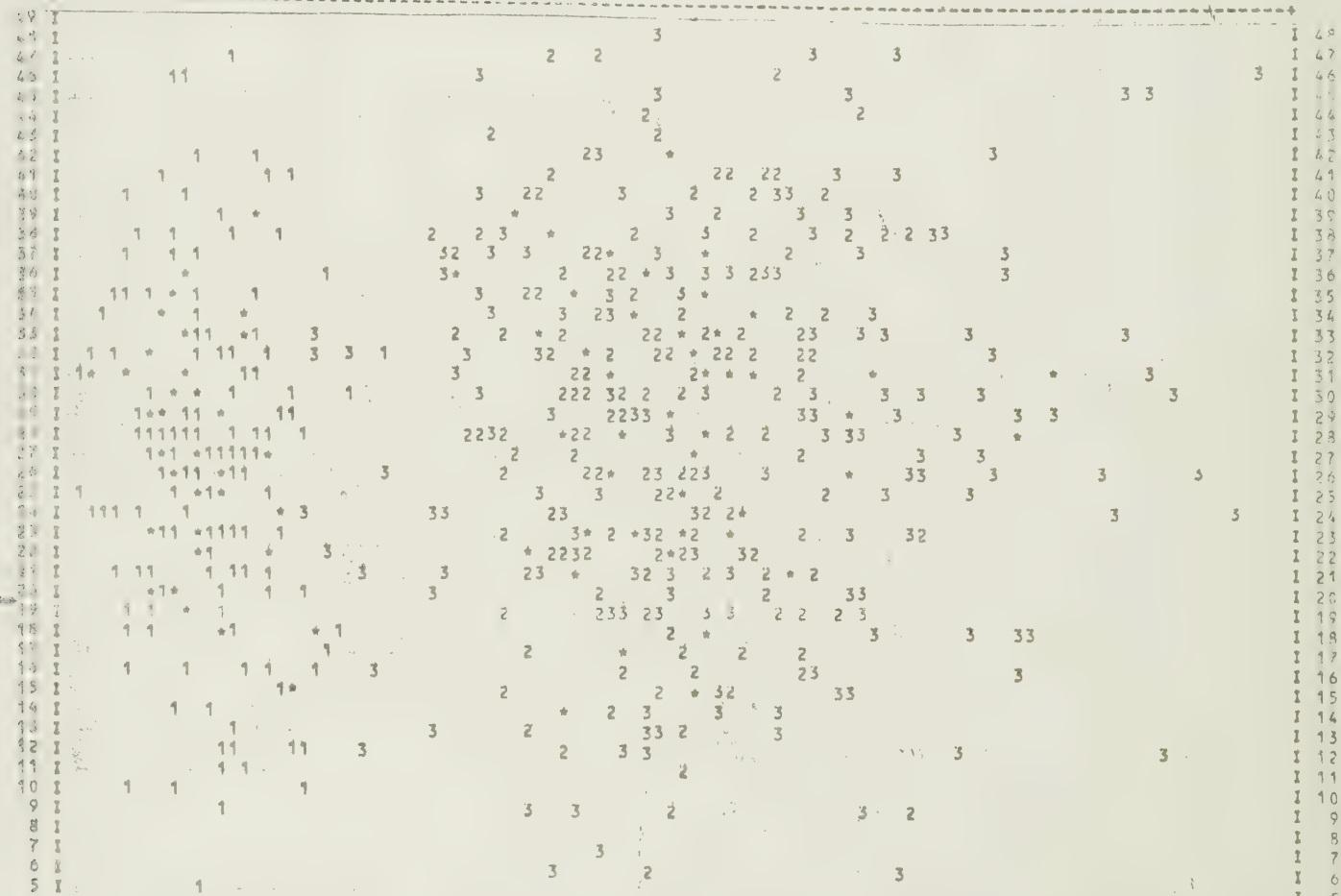
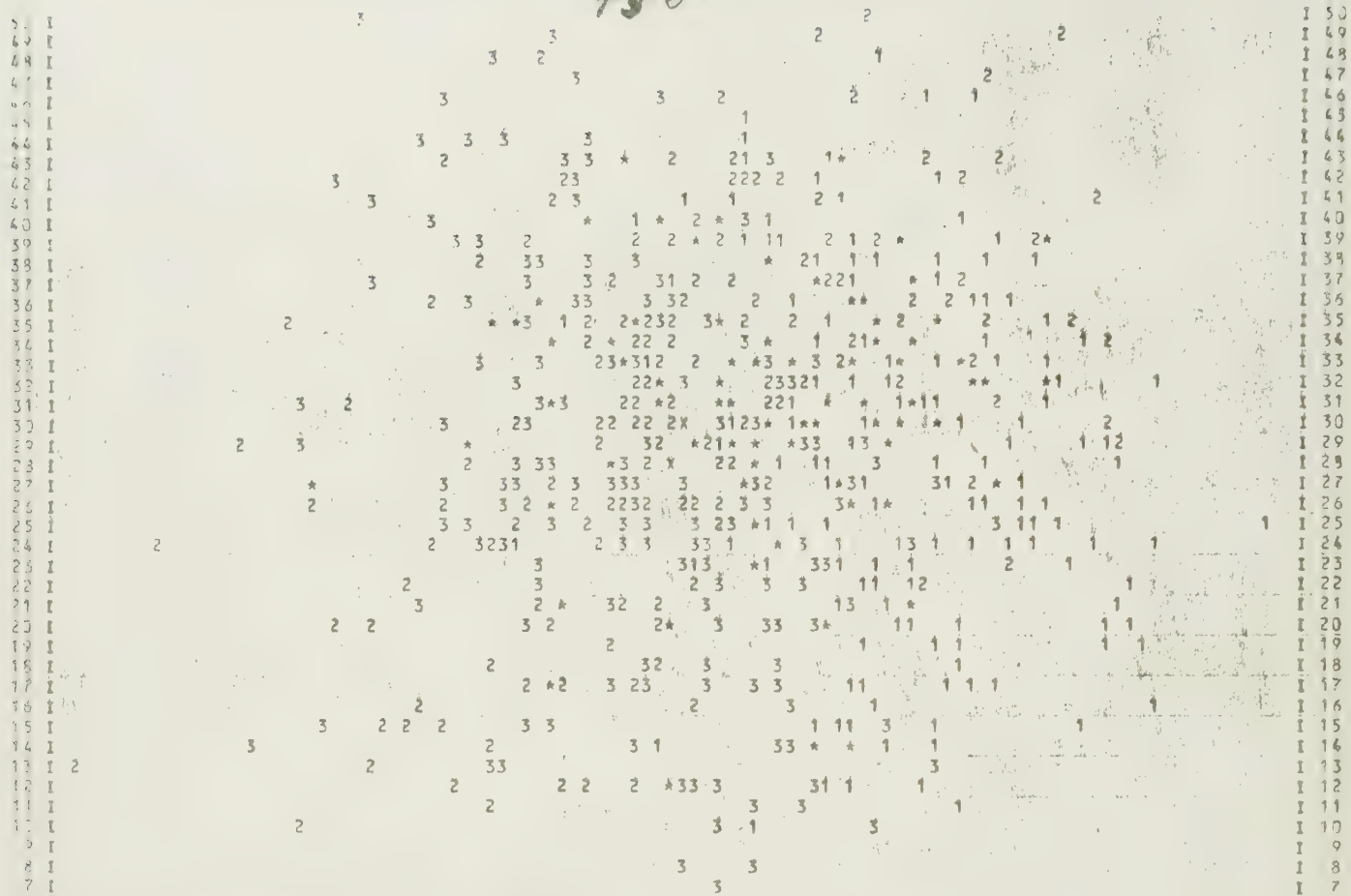


Abb.8 - Abbildung des Diskriminanzraumes

Oben - Reduktion von 10 auf 5 Dim.(Hauptachsentr.)

Unten - informationstheor. Redukt. von 10 auf 5 Dim.

theoretischen Reduktion von 10 auf 5 Dimensionen abgebildet. Die Verschiebung der ersten Population bleibt auch nach der Reduktion erhalten.

In diesem Fall werden bei Neuklassifikationen durchschnittlich 30% Fehler ohne und mit Interaktionen gefunden. Die immer noch relativ hohe Fehlerquote ist dadurch zu erklären, daß solche Datenvektoren neuklassifiziert wurden, die aus einem Datenpool von zwei konzentrischen Populationen und von einer verschobenen Population stammen (s. Tabelle 5 u. 6).

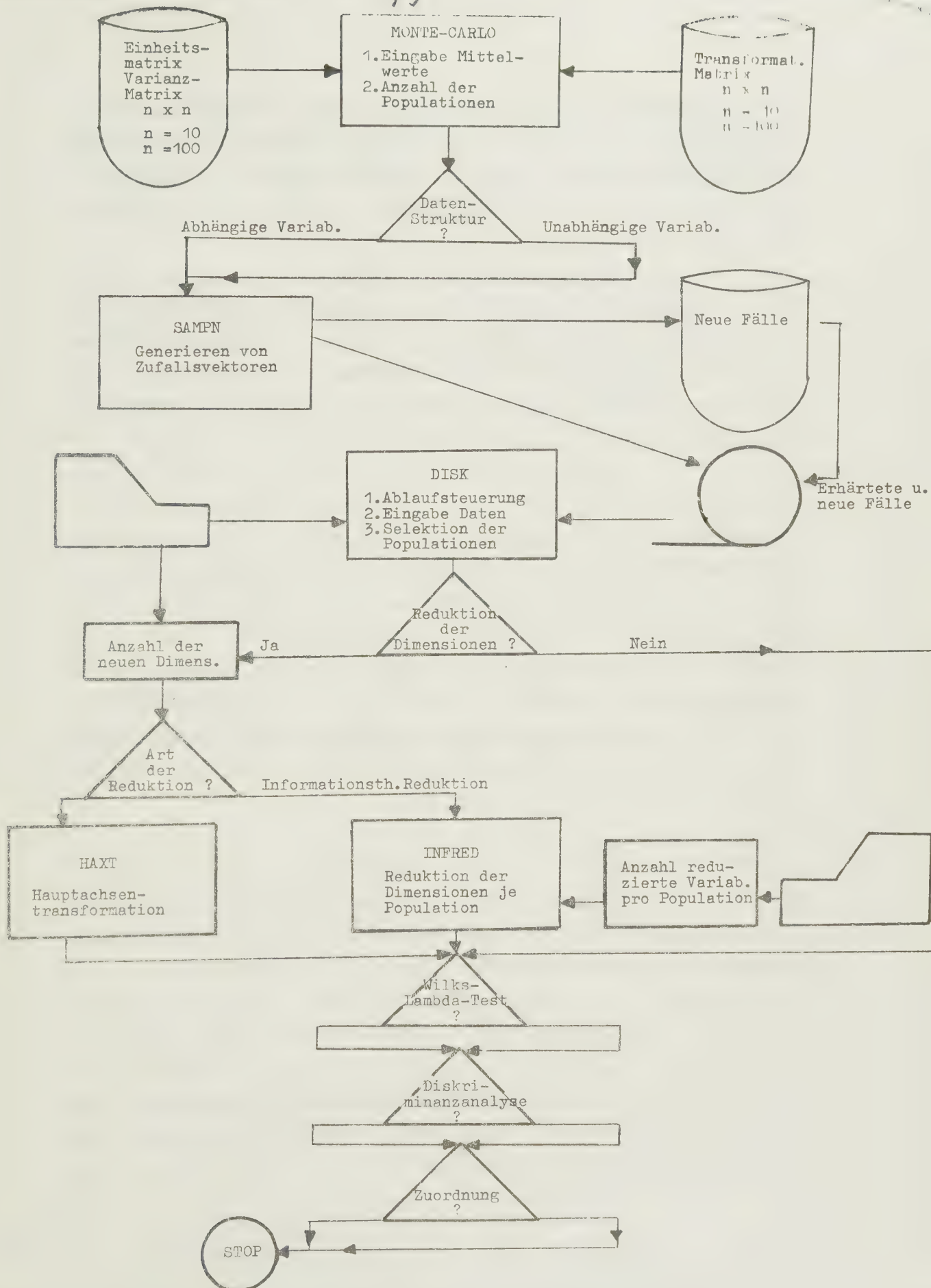
Der Anhang C zeigt Ergebnisse des DISK- Programms. Dabei wurde zur Reduktion der Dimensionen einmal die Hauptachsentransformation und zum anderen das Prinzip der maximalen Variablendivergenz verwendet.

8.3. Programmaufbau

Das Simulationssystem besteht aus einem Datenzufallsgenerator, einer Datenbank und einem Programm für Diskriminanzanalyse und Klassifikation (s. Abb. 9).

Das Hauptprogramm MONTE-CARLO und zwei Unterprogramme SAMPN und PROYE bilden den Datenzufallsgenerator. Das Programm MONTE-CARLO dient zur Eingabe der Modellparameter, die teilweise im voraus auf feste Dateien gespeichert werden (Varianzmatrix, Transformationsmatrix), oder vom Datensichtgerät (eventuell Lochkarten) in Dialogform eingegeben werden können (Mittelwerte, Anzahl der Populationen).

Mit Hilfe des Unterprogramms PROYE kann die Datenstruktur geändert werden, und zwar durch die Projektion der Daten



in einen Raum, in dem Abhängigkeiten zwischen den Variablen vorliegen. Die Transformationsmatrix bestimmt dann die Abhängigkeiten und dient auch für die Berechnung der neuen Varianz-Kovarianz-Matrix, bzw. des neuen Mittelpunktsvektors.

Das zweite Unterprogramm des Datenzufallsgenerators ist mit SAMPN gekennzeichnet und erzeugt pseudozufällige m-dimensionalen Vektoren mit voneinander unabhängigen Komponenten. SAMPN muß für jeden Vektor erneut aufgerufen werden.

Im Unterprogramm SAMPN wird jede Komponente des m-dimensionalen Zufallsvektors unabhängig mit Hilfe der Subroutine NORMAL (Programmbibliothek des Rechners IRIS 80) erzeugt. Für jede Population werden beim ersten Aufruf des Unterprogramms SAMPN die Anlaufszahlen dieser einzelner Zufallsgeneratoren als gleichverteilte Zahlen im Intervall (0,1) erstellt. Dafür hilft die Subroutine ALEAT des standardisierten Programmpaketes des Rechners IRIS 80 bei der Erzeugung pseudozufälliger, im Intervall (0,1) gleichverteilter Zahlen.

ALEAT liefert in (0,1) gleichverteilte Zufallsgrößen, die nach der additiven Kongruenzmethode erzeugt werden. Zunächst wird eine im Intervall $(0, 2^{31})$ gleichverteilte Zufallszahl nach folgendem Verfahren erstellt:

$$z_{i+1} = a z_i + c \quad (\text{modulo } m) \quad ,$$

wobei für a und m folgende optimale (bestimmen eine maximale Periodizität für einen Rechner mit 32 Bits pro Wort) Werte angenommen werden:

$$a = 2^{16} + 3 \quad \text{und} \quad m = 2^{31} \quad .$$

Die erzeugten Zahlen haben in diesem Fall die Periode

$$h = 2^{29} \quad .$$

Als Anfangszahl wird $c = 101$ verwendet. Um weiter gleichverteilte Zufallszahlen im Intervall $(0,1)$ zu erhalten werden die Zahlen z_i mit $1/m$ multipliziert.

Das Unterprogramm NORMAL erzeugt normalverteilte Zufallszahlen nach vorgegebenen Erwartungswert und Varianz. Folgendes Prinzip (s.Theorem (4.2.3.))[2]) wird dabei zu Grunde gelegt : Die Summe von N im Intervall $(0,1)$ gleichverteilte Zufallszahlen z_i (mit Erwartungswert $1/2$ und Varianz $1/12$) ergibt für ausreichend großen N eine normalverteilte normierte Zufallszahl:

$$\frac{\sum_{i=1}^N z_i - \frac{N}{2}}{\frac{N}{12}} = \frac{x - \mu}{\sigma}$$

Eine normalverteilte Zufallszahl wird also durch Addition von N im Intervall $(0,1)$ gleichverteilte Zufallszahlen nach folgendem Verfahren erzeugt:

$$x = \mu + \sigma \frac{12}{N} \left[\sum_{i=1}^N z_i - \frac{N}{2} \right] ,$$

wobei $N = 12$ ist.

Die Datenbank besteht aus Datensätze mit unterschiedlichen Stichproben. Sie ist in Dateien eingeteilt, wobei eine Datei eine bestimmte Anzahl der Probanden (150, 300 und 900) pro Datensatz enthält. Jede Datei enthält vier Datensätze, drei davon bestehen aus erhärteten Fällen der jeweiligen Population, während der vierte aus neuen Fällen (eine Mischung aus allen Populationen) gebildet wird. Dazu gehört auch eine Datei, die Populationen mit 60 Variablen enthält.

Die aufgebauten Dateien werden mit Hilfe von statistischen Programmen verarbeitet. Das Programm DISK für multiple Diskriminanzanalyse und Zuordnungsverfahren aus dem ehemaligen DRZ in Darmstadt, wird zu diesem Zweck verwendet, wobei

die Unterprogramme INFRED und ZWEIP zur informationstheoretischen Reduktion der Dimensionen, von mir geschrieben und implementiert wurden. Damit ist es möglich, eine Reduktion der Dimensionen, eine Diskriminanzanalyse, einen Wilks'-Lambda-Test zur Prüfung von Unterschieden zwischen den Gruppen und schließlich Klassifikationen durchzuführen.

9. Schlußbemerkung

In der vorliegenden Arbeit wurde die Selektion bestimmter Merkmale bei der Diagnosefindung, vom informationstheoretischen Gesichtspunkt betrachtet und mit Hilfe von simulierten Patientendaten praktisch durchgeführt. Nach der Reduktion der Dimensionen (Merkmale) wurde eine minimale Fehlklassifikationsrate erreicht.

Die Hauptkomponentenmethode wird unter informationstheoretischen Betrachtung diskutiert und mit Hilfe von simulierten Daten überprüft. Die daraus resultierende Ansätze heben eine wichtige Einschränkung dieser Methode hervor: die Dimensionenreduktion wird nur aufgrund der in den Varianzen enthaltenen Trenninformation durchgeführt. Der Teil der gesamten Trenninformation, der in den Mittelpunktsvektoren enthalten ist, wird dabei nicht berücksichtigt.

Andere Methoden zur Reduktion der Dimensionen, die im Kapitel 6. beschrieben sind, stützen sich auf das Testen der Nullhypothese. Aufgrund des Tests auf Gleichheit der Mittelpunktsvektoren (Nullhypothese) wird hier ein Maß der Diskriminanzleistung jeder Variablen abgeleitet. Dieses Maß beschreibt nur die in den Mittelpunktsvektoren enthaltenen Trenninformation, wie es aus der Betrachtung der Test-Statistik [Kap.6., Formel (6.6)] folgt. Es ist unmittelbar daraus zu schließen, daß alle Methoden, die sich auf den oben beschriebenen Test stützen, einen wichtigen Teil der Trenninformation vernachlässigen, und zwar diejenige Trenninformation, die in den Varianzen enthalten ist. Es liegen jedoch keinerlei experimentelle Überprüfungen dieser Behauptung vor. Diese könnten eventuell Gegenstand einer zukünftigen Arbeit sein.

Das Hauptziel dieser Arbeit ist, ein Kriterium zu finden, das bei der Reduktion der Dimensionen die gesamte, in den Variablen (Markmalen) enthaltene Trenninformation berücksichtigt. Es wurde deshalb theoretisch und experimentell bewiesen, daß die Variablendivergenz ein optimales Trenninformationsmaß darstellt und dadurch für eine Reduktion der Dimensionen am besten geeignet ist.

Literaturverzeichnis

1. Aitchison, J. and Kay, J.W. "Principles, practice and performance in decision-making in clinical medicine", 1973 Nato Conference, Luxembourg
2. Anderson, T.W., "An Introduction to Multivariate Statistical Analysis", 1958, J.Wiley, New York
3. Andrews, D.F., "Plots of High-Dimensional Data", Biometrics, March 1972, S. 125
4. Bahadur, R.R., "A representation of the joint distribution of responses to n dichotomous items", Studies in Item Analysis and Prediction, Stanford University Press, California, 1961, S.1586
5. Bailey, N.T.J., "Probability methods of diagnosis based on small Samples", Biology and Medicine, Her Majesty's Stationary Office, 1965, London
6. Balanescu, A., "Ärztliche Diagnose unter Anwendung der elektronischen Rechenmaschinen", Automatica si Electronica , Nr.6 , Bukarest, 1973
7. Barnard, G.A., "Sampling Inspection and Statistical Decisions", Journal of the Royal Statistical Society, Series B, Vol. 16, 1954, pp. 151-174
8. Birch, M.W., "Maximum Likelihood in Three-Way Contingency Tables", Journal of the Royal Statistical Society, Series B, Vol.25, 1963, p.220
9. Blomer, R.J. und Drechsler K.H., "Vorgehen bei der Klassifizierung und Kreuzvalidisierung großer Datenmengen", 19.Jahrestagung der DGMS, Mainz 1974
10. Bouckaert, A., "Computer-aided diagnosis of goitres in a cancer department" International Journal of Bio-Medical Computing, nr. 3, 1972, p.3

11. Bouckaert, A. and Beckers, C., "Computer-aided differential diagnosis of goitres in internal medicine", International Journal of Bio-Medical Computing, nr.4, 1973, p.1
12. Christl, H.L., und Stock, S., "Untersuchung des Einflusses von systematischen und zufälligen Fehlern bei der Symptomerhebung auf die Leistungsfähigkeit automatisierter Verfahren zur Diagnoseunterstützung", Bericht MD 95 der GSF München, 1974
13. Cooley, W.W. and Lohnes, P.R., "Multivariate Data Analysis", J.Wiley, 1971, New York
14. Cox, D.R., "The Analysis of Binary Data", Menthuen London, 1970
15. Cumberbatch, J. and Heaps, H.S., "Application of a non-bayesian approach to computer-aided diagnosis for upper abdominal pain", International Journal of Bio-Medical Computing, nr.4, 1973, p.105
16. Cumberbatch, J. and Heaps, H.S., "A non-probabilistic method for automatic medical diagnosis", Int. J. of Bio-Med. Comp., nr.5, 1974, p.133
17. Dixon, W.J., "BMD - Biomedical Computer Programs", University of California Press, 1970
18. Fischer Lexikon, Medizin 1 und 2, 1959, S.169
19. Fisher, R.A., "The statistical utilization of multiple measurements", Annals of Human Genetics, Vol.8, 1937-38, p.376
20. Fisher, R.A., "The use of multiple measurements in taxonomic problems", Annals of Human Genetics, Vol.7, 1936-37
21. Fisher, R.A., "Theory of statistical estimation", Proc. Cambrige Philos. Soc., Vol. 22, 1925, p.700
22. Geisser, S., "Posterior odds for multivariate normal classifications", J.of the Roy. Stat. Soc., B, nr.26, 1964, p.69

36. Lasker, G.E., "Application of sequential pattern-recognition technique to medical diagnostics", Int. J. of Bio-Med. Comp., nr.1, 1970, p.473
37. Ledley, R.S. and Lusted, L.B., "Reasoning Foundations of Medical Diagnosis", Science, 130, 1959
38. Martin, D.C. and Bradley, R.-A., "Probability models, estimation, and classification of multivariate dichotomous populations", Biometrics, nr.28, 1972, p.203
39. Morrison, D.F., "Multivariate Statistical Methods", Mc Graw-Hill, New York, 1967
40. Okamoto, M., "An asymptotic expansion for the distribution of the linear discriminant function", Annals of Math. Stat., nr.34, 1963, p.1286
41. Press, S.J., "Applied Multivariate Analysis", 1972 Holt, Rinehart and Winston Inc., 1971
42. Rao, C.R., "Advanced statistical methods in biometric research", J.Wiley, New York 1967
43. Reale, A. et al. "Computer diagnosis of congenital heart disease", Comput. Biomed. Research, nr.1, 1968, p.533
44. Rinaldo, J.A., Scheinok, P. and Rupe, C.E., "Symptom diagnosis: A mathematical analysis of epigastric pain", Annals of Internal Medicine, nr.59, 1963, p.145
45. Sammon, J.W., "A nonlinear Mapping for Data Structure Analysis", IEEE Trans. on Comp., May 1969
46. Scheinok, P., "Symptom diagnosis: Bayes's Theorem and Bahadur's Distribution", J. of Bio-Med. Comp., nr.3, 1972
47. Scheinok, P. and Rinaldo, J.A., "Symptom diagnosis: Optimal subsets for upper abdominal pain", Comput. Biomed. Research, nr.1, 1967, p.221
48. Scheinok, P. and Rinaldo, J.A., "Symptom diagnosis: A comparison of mathematical models related to upper abdominal pain", Comput. Biomed. Research, nr.1, 1968, p.457

23. Gilbert, E.S., "On discrimination using qualitative Data", J. Am. Stat. Assoc. , nr.63, 1968, p.1399
24. Gilbert, E.S., "The Effect of unequal Variance-Covariance-Matrices on Fisher's Linear Discriminant Function", Biometrics, nr.25, 1969, p.505
25. Good, I.J., "Maximum Entropy for Hypothesis Formulation, Especially for Multidimensional Contingency Tables", Annals of Mat. Statistics, nr. 34, 1963, p.911
26. Gorry, G.A., "Strategies for computer-aided diagnosis", Math. Biosciences, nr. 2, 1968, p.293
27. Gorry, G.A. and Barnet, G.O., "Experience with a model of sequential diagnosis", Comput. Biomed. Research, nr.1, 1968, p.490
28. Hills, M., "Allocation Rules and their Error Rates", J. of the Royal Stat. Soc., B, nr.28, 1966, p.1
29. Hölzel, D., "Ein Informationssystem zur Diagnoseunterstützung", Dissertation, Abt. für Med. Statistik, Dokumentation und Datenverarbeitung der Universität Ulm, 1973
30. Jeffreys, H., "Theory of Probability", Oxford, 1961
31. Kendall, M., "Multivariate Analysis", Charles Griffin, London, 1975, p.159
32. Krischer, J.P., "Diagnostic Nets", Int. J. of Bio-Med. Comp., nr.2, 1971, p.27
33. Kullback, S., "Application of Information Theory to Multivariate Analysis", Annals of Math. Stat., 1952, p.88
34. Kullback, S. and Leibler, R.A., "On information and sufficiency", Annals of Math. Stat., Vol.22, 1951, p.79
35. Lachenbruch, P.A., "Estimation of error rates in discriminant analysis", Ph. D. Dissertation, Univ. of California, 1965, Los Angeles

49. Scheinok, P. and Rinaldo, J.A., "Symptom diagnosis: The use of two different mathematical models", Int. J. of Bio-Med. Comp., nr.2, 1971, p.239
50. Stein, M.A. and Winter, J., "Theory Development in Medical Diagnosis", Int.J. of Bio-Med. Comput., nr5 , 1974
51. Warner, H.R., Toronto, A.F., Veasey, L.G. and Stephenson, R., "A mathematical approach to medical diagnosis: Application to congenital heart disease", J. Am. Med. Association, nr.177, 1961, p.177
52. Wilks, S.S., "Certain generalisations in the Analysis of Variance", Biometrika, nr.24, 1932, p.471

Im folgenden soll auch die Literatur, die nicht ausdrücklich zitiert wurde, aber auf die Entstehung der hier referierten Gedanken wesentlichen Einfluß hatte, genannt werden:

Anderson, T.W. and Bahadur, R.R., "Classification into two multivariate normal distributions with different covariance matrices", Annals of Math. Stat., nr.33, 1962 p.420

Blomer, R.J. und Drechsler, K.H., "Hold-one-out probabilities for multivariate normal classification", IMD der GSF - München, 1975

Cochran, W.G., "On the performance of the linear discriminant function", Int.Stat.Institute-Bulletin, nr.39, 1962, p.435

- Cochran, W.G. and Hopkins, C.E., "Some classification problems with multivariate qualitative data", *Biometrics*, nr. 17, p. 10
- Cox, D.R., "The Analysis of Binary Data", London, Menthuen, 1970
- Dadaschew, R.S., "Waprosi avtomatitscheskogo diagnosa sostoianii organisma", Moskau, Nauka, 1967
- Elashoff, J.D., Elashoff, R.M. and Goldman, G.E., "On the choice of variables in classification problems with dichotomous variables", *Biometrika*, nr. 54, 1967, p. 668
- Geisser, S. and Cornfield, J., "Posterior Distributions for Multivariate Normal Parameters", *J.R. Stat. Soc., B*, nr. 25, p. 368
- John, S. "Errors in Discrimination", *Annals of Math. Stat.*, nr. 32, 1961, p. 1125
- Kendall, M.G. and Stuart, A., "The Advanced Theory of Statistics", 3 Vol., Charles Griffin, London, 1958
- Mantel, N., "Models for Complex Contingency Tables and Polychotomous Dosage Response Curve", *Biometrics*, nr. 22, p. 83
- Nicolau, E. si Popovici, A., "Algoritmi, automate finite, calculatoare electronice", Bukarest, Ed. stiintifica, 1970
- Solomon, H. "Studies in Item Analysis and Prediction", Stanford Univ. Press, 1961
- Uesaka, Y. "Construction of a complex-valued nonlinear discriminant function for pattern recognition", *IEEE Tr. on Systems, Man and Cyb.*, nr. 3, 1971, p. 194
- Zehna, P., "Probability Distributions and Statistics", Allyn and Bacon Inc., Boston, 1970

ANHANG A

Parameterschätzung für normalverteilte Populationen

1. Maximum-Likelihood-Schätzung

Wir untersuchen zunächst das Problem, wie man eine Stichprobe $\mathbf{z} = \{\mathbf{x}^i, (i = 1, 2, \dots, N)\}$ von m -dimensionalen unabhängigen Probanden aus der Population $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ zieht. Die Annahme der statistischen Unabhängigkeit zwischen den Probanden $\mathbf{x}^i$ gibt die Möglichkeit, die gemeinsame Dichtefunktion der Vektoren $\mathbf{x}^i, (i = 1, 2, \dots, N)$ durch Multiplikation der entsprechenden Randdichtefunktionen zu erhalten:

$$(a.1) \quad f(\mathbf{z}/\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{Nm}{2}} |\boldsymbol{\Sigma}|^{\frac{N}{2}}} \exp\left[-\frac{1}{2} \sum_{i=1}^N (\mathbf{x}^i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}^i - \boldsymbol{\mu})\right].$$

Diese Funktion ist als Likelihood-Funktion der Stichprobe $\mathbf{z}$ oder als Sample-Dichtefunktion bekannt. Geht man von der Annahme aus, daß die bedingte Wahrscheinlichkeit $f(\mathbf{z}/\boldsymbol{\mu}, \boldsymbol{\Sigma})$ - eine Stichprobe $\mathbf{z}$ von der Population $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ zu erhalten - von den optimalen Schätzungen $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}$ maximiert wird, so werden durch aufeinanderfolgenden Ableitungen von (a.1) die Schätzungen

$$(a.2) \quad \begin{aligned} \hat{\boldsymbol{\mu}}_{ML} &= \frac{1}{N} \sum_{i=1}^N \mathbf{x}^i = \bar{\mathbf{x}} \\ \hat{\boldsymbol{\Sigma}}_{ML} &= \frac{1}{N} \sum_{i=1}^N (\mathbf{x}^i - \bar{\mathbf{x}})(\mathbf{x}^i - \bar{\mathbf{x}})^T \end{aligned}$$

erhalten, die unter dem Namen Maximum-Likelihood-Schätzungen bekannt sind.

Da diese Schätzungen von der Größe der Stichprobe $\mathbf{z}$ abhängig sind, können sie selbst als Zufallsgrößen betrachtet werden. Unter dieser Voraussetzung genügt $\bar{\mathbf{x}}$ einer

Normalverteilung([2],S.53,159)

$$(a.3) \quad N(\mu, \frac{\Sigma}{N})$$

und einer Wishartverteilung

$$(a.4) \quad W(\Sigma, m, n=N-1) \quad .$$

2. Bayes-Schätzung

Falls keine apriori Informationen über die Verteilungsparameter vorliegen (diffuser Fall) , können jedoch einige Beschlüsse über die Struktur dieser Parameter gemacht werden. Jeffreys [30](1961) hat auf der Suche nach Invarianten beim Parameterwechsel eines Schätzungsproblems herausgefunden, daß eine apriori Verteilung der Parameter trotz der fehlenden apriori Information folgende Bedingung erfüllen muß

$$(a.5) \quad f(\mu, \Sigma) \propto \frac{1}{|\Sigma|^{\frac{m+1}{2}}} \quad .$$

Eine andere Art die geringe apriori Information auszudrücken, stammt von Barnard [7](1954) und ist mittels der natürlich konjugierten apriori Verteilung durchzuführen. Bei natürlich konjugierte apriori Verteilung entsteht aus einem Rollentausch zwischen Zufallsvariablen und Parametern. Die ersten bleiben fix und vorgegeben erhalten und die letzteren werden als zufällig betrachtet. Wenn die neue Verteilung durch Änderung der Proportionalitätskonstante zur Einheit integriert wird und die Parameter als veränderlich betrachtet werden (nicht auf den geläufigen Daten), findet man eine Dichtefunktion, die in diesem Sinn, natürlich konjugiert mit der ursprünglichen Dichtefunktion ist.

Aufgrund dieser Bemerkungen ist daraus zu schließen, daß

die gemeinsame natürlich konjugierteapriori Verteilung für μ und $\Lambda = \Sigma^{-1}$, unter der Bedingung $n > 2m$, durch die folgende normale Wishartverteilung gegeben ist [41]

$$(a.6) \quad f(\mu, \Lambda) \propto |\Lambda|^{\frac{p-m}{2}} \exp\left[-\frac{1}{2}(\mu - \alpha)^T \Lambda (\mu - \alpha)b + \text{tr} \Lambda G^{-1}\right].$$

Dabei genügt Λ einer Randverteilung $W(G, m, p)$, d.h. einer normalen Wishartverteilung mit Maßstabmatrix G , m Dimensionen und p Freiheitsgraden. Die Matrix G und die Freiheitsgrade p können willkürlich ausgewählt werden mit der Einschränkung, daß $G > 0$ und $p \geq m$ ist. Die Parameter G, p, α und b können nach subjektiven Kriterien, allgemein bekannten Tatsachen, Erfahrung etc. ausgewählt werden.

Da bekanntlich die Größen $\bar{x}$ und S unabhängig voneinander sind, wird ihre gemeinsame Dichtefunktion durch Multiplikation der Randdichtefunktionen entsprechend der Verteilungen (a.3) und (a.4) erhalten

$$(a.7) \quad f(\bar{x}, S / \mu, \Sigma) \propto \frac{|S|^{\frac{n-m-1}{2}}}{|\Sigma|^{\frac{N}{2}}} \exp\left\{-\frac{1}{2}\text{tr} \Sigma^{-1} [S + N(\bar{x} - \mu)(\bar{x} - \mu)^T]\right\},$$

wobei $n = N-1$. Unter der Anwendung der diffusen apriori Verteilung (a.5) kann dann die aposteriori Verteilung der Parameter abgeleitet werden:

$$(a.8) \quad f(\mu, \Sigma / \bar{x}, S) \propto \frac{1}{|\Sigma|^{\frac{N+m+1}{2}}} \exp\left\{-\frac{1}{2}\text{tr} \Sigma^{-1} [S + N(\bar{x} - \mu)(\bar{x} - \mu)^T]\right\}.$$

Berechnet man durch Integration die Randverteilungen $f(\mu / \bar{x}, S)$ bzw. $f(\Sigma / \bar{x}, S)$ und danach ihre Momente ersten Grades, so ergeben sich folgende Bayes-Schätzungen:

$$(a.9) \quad \begin{aligned} \hat{\mu}_{BD} &= \bar{x} \\ \hat{\Sigma}_{BD} &= \frac{1}{N-m-2} S \quad \text{wenn } N > 1 \end{aligned}$$

Wiederholt man diese Prozedur mit der natürlich konjugierten apriori Verteilung (a.6) , so erhält man folgende Schätzungen

$$(a.10) \quad \begin{aligned} \mu_B &= \frac{\alpha + N \bar{x}}{1 + N} \quad , \quad N+p-m > 1 \\ \Sigma_B &= \frac{S + G + \frac{N}{1+N} (\bar{x} - \mu)(\bar{x} - \mu)^T}{N + p - 2m - 2} \quad , \quad N+p-2m > 2 . \end{aligned}$$

Unter der Annahme folgender Vereinfachungen

$$\alpha = aI \quad , \quad G = \begin{pmatrix} h & 1 & \dots & 1 \\ 1 & h & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & h \end{pmatrix} \quad \text{mit} \quad 1 \geq - \frac{h}{m-1} \quad ,$$

müssen insgesamt nur 5 Parameter (p,b,a,h,l) willkürlich ausgewählt werden.

ANHANG B

Testen der Nullhypothese

Unter der Voraussetzung einer Dichtefunktion $f(x/\theta)$ mit $\theta \in \Omega = \{\Omega_0, \Omega_1\}$, wobei $\Omega, \Omega_0, \Omega_1$ den Parameter-raum bzw. zwei Unterräume des Parameterraumes darstellen, kann eine Hypothese H_0 bzw. ihre Alternative H_1 wie im folgenden definiert werden

$$(b.1) \quad \begin{aligned} H_0 &: \theta \in \Omega_0 \\ H_1 &: \theta \in \Omega_1 \end{aligned}$$

Die Hypothese H_0 teilt den Stichprobenraum in zwei Regionen ein: in eine Region S_0 , die der Annahme der Hypothese H_0 entspricht und in eine Region S_1 , die der Zurückweisung dieser Hypothese entspricht (oder der Annahme der Hypothese H_1 entspricht) (Abb. 15). Dabei wird eine Entscheidungsfunktion (Test) $d: S \rightarrow A$ als Korrespondenz zwischen dem Stichprobenraum S und dem Aktionsraum $A = \{a_0, a_1\}$ definiert. Je nachdem die Entscheidung für H_0 oder für H_1 fällt, ist für jede Realisation $x \in S$ das Bayes'sche Risiko

$$(b.2) \quad r = \begin{cases} \alpha(\theta) = P_\theta[x \in S_1] & \text{- Wahrscheinlichkeit für } x \in S_1 \\ & \text{(wenn } \theta \text{ der wahre Parametervektor ist und } \theta \in \Omega_0) \\ \beta(\theta) = 1 - \alpha(\theta) = \\ = P_\theta[x \in S_0] & \text{- Wahrscheinlichkeit für } x \in S_0 \\ & \text{(wenn } \theta \text{ der wahre Parametervektor ist und } \theta \in \Omega_1) \end{cases}$$

zu berücksichtigen. Der Wert

$$\tau = \max \{ \alpha(\theta), \theta \in \Omega_0 \}$$

wird im weiteren als Signifikanzpegel des Hypothesentests bezeichnet. Die Leistungsfunktion des Tests schreibt sich dann

$$(b.3) \quad \pi(\theta) = \begin{cases} \alpha(\theta) & \text{für } \theta \in \Omega_0 \\ 1 - \alpha(\theta) & \text{für } \theta \in \Omega_1 \end{cases} .$$

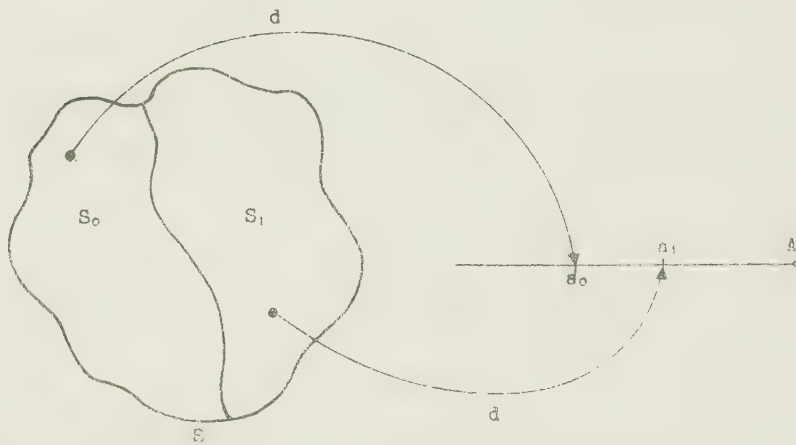


Abb. 10 - Abbildung des Stichprobenraumes in den Aktionsraum
 $d : S \rightarrow A$

Gesucht ist der Test d mit Signifikanzpegel α [$\pi(\theta) \leq \alpha$ für alle $\theta \in \Omega_0$], der die Leistungsfunktion π über alle Werte $\theta \in \Omega_1$ maximiert.

Konkreter: für eine Realisation $x \in S$ wird zunächst die Likelihood-Funktion L_x über alle Werte $\theta \in \Omega_0$ maximiert. Das Maximum wird mit $L_x(\hat{\theta}_0)$ bezeichnet und wenn Ω_0 nur aus einem Vektor θ_0 besteht ($\Omega_0 = \{\theta_0\}$) dann

$$L_x(\hat{\theta}_0) = L_x(\theta_0) .$$

Danach wird L_x über den gesamten Raum Ω maximiert. Das erreichte Maximum $L_x(\hat{\theta})$ ist aber die Maximum-Likelihood-Schätzung $\hat{\theta}$ (s. Anhang A) des Parametervektors θ . Daraus folgt

$$L_x(\hat{\theta}) = L_x(\hat{\theta}) .$$

Aus den maximalen Werten bildet man den Quotienten (der eine Funktion von $\mathbf{x}$ ist)

$$(b.4) \quad \Lambda_{\mathbf{x}} = \Lambda(\mathbf{x}) = \frac{L_{\mathbf{x}}(\theta_0)}{L(\hat{\theta})} \quad ,$$

der als verallgemeinerter Likelihood-Quotient bekannt ist.

Unter der Annahme einer Normalverteilung ist (b.4) ein Quotient aus zwei quadratischen Formen. Infolgedessen genügt die Statistik $\Lambda(\mathbf{x})$ einer F-Verteilung. Der Fehler $\alpha(\theta)$ schreibt sich dann

$$(b.5) \quad P_{\theta}[\mathbf{x} \in S_1] = P_{\theta}[\Lambda(\mathbf{x}) > k] = \alpha \quad ,$$

wobei α den Signifikanzpegel darstellt. Die Unbekannte der Gleichung (b.5) ist k . Von der Verteilung der Statistik $\Lambda(\mathbf{x})$ ist k durch den Zusammenhang

$$k = F_{n_2, \alpha}^{n_1}$$

zu bestimmen, wobei n_1 und n_2 die Freiheitsgrade des Zählers bzw. des Nenners bezeichnen.

Wenn q m -dimensionale normalverteilte Populationen vorliegen, kann die Nullhypothese

$$(b.6) \quad H_0 : \mu^1 = \mu^2 = \dots = \mu^q$$

mit Hilfe der Wilks' F-Quotienten-Teststatistik (Λ) (s. §6.3) gegen die Alternative

$$H_1 : \mu^k \neq \mu^p \quad \text{für } k \neq p$$

getestet werden. Wilks' Lambda-Quotient

$$\Lambda = \frac{|\Sigma|}{|\mathbf{T}|}$$

genügt einer Verteilung, die mit Hilfe der $F_{n_2}^{n_1}$ Dichte
angenhähert werden kann (s. §6.3). Ist diese Dichtefunktion
durch den Zusammenhang (6.7) gegeben, so wird die Hypo-
these H_0 nur dann mit Konfidenz $1 - \alpha$ angenommen
werden, wenn

$$(b.7) \quad \Lambda^{1/S} > F_{n_2, \alpha}^{n_1}$$

ist. Dieser Test wurde aufgrund von q Stichproben
mit der gleichen Sample-Varianz-Kovarianz-Matrix S aus
den berücksichtigten Populationen aufgebaut. Er kann erst
dann durchgeführt werden, wenn sich ein Test auf gleichen
Varianz-Kovarianz-Matrizen als erfolgreich erwiesen hat
([13], S.228).

ANHANG C

Ergebnisse der Klassifikationsprogramme

DIMENSIONSVERKLEINERUNG
+++++

ANZAHL DER VARIABlen 3
ANZAHL DER VARIABLEN 10
ANZAHL DER OBJEKTE 600

KREUZPRODUKTMATRIX

SIGM SECTION 1

ROW	1	2	3	4	5	6	7	8	9	10
1	909.7920	259.6119	15.6474	1.5221	268.7396	59.5066	218.2055	303.6987	42.0037	-3.6798
2	250.6119	947.4783	63.5429	-15.5767	76.6782	101.0945	103.9911	76.9964	-12.4068	-21.2481
3	295.0424	63.5429	181.3741	-17.1804	66.8127	59.6014	57.7809	75.3230	23.0299	-1.6353
4	1.5221	-15.5767	-17.1804	229.7013	-13.1075	25.3324	2.5248	12.5356	5.6772	-6.7781
5	268.7396	76.6782	66.8127	-13.1075	188.4140	-26.9243	11.1939	84.4362	22.9190	-16.8255
6	59.5066	101.0945	59.6014	25.3324	-26.8243	822.6228	-19.6423	60.2803	20.3508	34.5224
7	218.2055	103.9911	57.7809	2.5248	11.1939	-19.6423	713.8386	128.6370	9.4758	-4.9292
8	303.6987	76.9964	75.3230	12.5356	84.4362	60.2803	128.6370	205.9060	12.4903	4.3590
9	42.0037	-12.4068	23.0299	5.6772	22.9190	20.3508	9.4758	12.4903	422.0955	-13.4963
10	-3.6798	-21.2481	-1.6353	-6.7781	-16.8255	34.5224	-4.9292	4.3590	-13.4963	79.3482

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6	7	8	9	10
1	164.1582	46.2348	36.7313	19.6823	45.8910	46.2758	30.1431	61.8657	.1002	16.0041
2	46.2348	859.9255	50.0569	14.0576	49.4983	14.8718	20.8656	56.3295	39.2889	14.0530
3	26.7313	50.0569	290.5477	13.1314	48.1598	48.7026	17.7431	68.1262	29.8011	30.1155
4	49.6823	14.0576	13.1314	144.4637	21.9307	7.0815	42.4062	32.5710	19.5299	14.1498
5	45.8910	49.4983	48.1598	21.9307	220.3975	105.8814	78.9480	90.2747	42.4635	58.8805
6	46.2758	14.8718	48.7026	7.0815	608.2817	77.2723	77.4073	80.3865	27.6148	-12.5640
7	30.1431	20.8656	17.7431	42.4062	78.9480	90.2747	601.4041	190.6348	8.7000	39.0240
8	61.8657	56.3295	68.1262	32.5710	90.2747	80.3865	90.6348	199.9664	13.5741	65.3822
9	.1002	39.2889	29.8011	19.5299	-2.4635	27.6148	8.7000	13.5741	516.9590	9.1098
10	16.0041	14.0530	30.1155	14.1498	58.8805	-12.5640	39.0240	65.3822	9.1098	443.9250

ALLGEMEINES EIGENWERTPROBLEM Y=V*W*V*W

SPUR DER MATRIX 14.12059

EIGENWERTE 5.8808 1.8109 1.6134 1.3120 1.0027 .8046 .7255 .3250 .1488

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

.30135E+02 .84485E+01 .74750E+04 -.14528E+01 .89120E+01 .19801E+01 .73365E+01 .10014E+02 .14336E+01 -.12106E+00
.59496E+01 .51313E+01 .67420E+04 .25344E+01 .83887E+01 .74041E+01 .92611E+01 .11206E+02 .30125E+01 .50363E+01
EIGENVEKTORENMATRIX FV(SPALTEN)

NO.	1	2	3	4	5	6	7	8	9	10
1	.1771	.0040	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2	.0004	.0040	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
3	.0013	.0013	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
4	.0143	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

DIVERGENZEN 1 1.4304E+01 1.8675E+00 1.247E+00 5.889E-01 1.173E-01 2.868E-01 1.208E+00 7.913E+00 9.410E+00 3.054E+01

DIVERGENZEN 1 1.714E+01 3.054E+01 1.247E+00 5.889E-01 1.173E-01 2.868E-01 1.208E+00 7.913E+00 9.410E+00 3.054E+01

EIGENVEKTORENMATRIX EV SPALTE 1

EV-M SECTION 1	1	2	3	4	5	6	7	8	9	10
ROW	1	2	3	4	5	6	7	8	9	10
1	.0774	.0077	.0177	.0183	.0151	.0028	.0120	.0029	.0036	.0032
2	.0004	.0016	.0010	.0014	.0040	.0090	.0123	.0010	.0057	.0302
3	.0143	.0050	.0463	.0183	.0013	.0030	.0154	.0010	.0025	.0004
4	.0069	.0035	.0049	.0012	.0533	.0585	.0037	.0291	.0045	.0108
5	.0092	.0148	.0062	.0082	.0298	.0315	.0223	.0292	.0057	.0171
6	.0016	.0016	.0006	.0006	.0306	.0218	.0099	.0128	.0005	.0049
7	.0034	.0017	.0008	.0091	.0068	.0157	.0161	.0340	.0037	.0041
8	.0018	.0048	.0145	.0178	.0045	.0089	.0185	.0127	.0320	.0307
9	.0018	.0047	.0009	.0004	.0082	.0029	.0121	.0022	.0397	.0155
10	.0039	.0421	.0147	.0173	.0053	.0062	.0064	.0028	.0005	.0016

KREUZPRODUKTMATRIX 2

SIGN SECTION 1	1	2	3	4	5	6	7	8	9	10
ROW	1	2	3	4	5	6	7	8	9	10
1	36.0715	10.6058	33.2046	13.3382	30.5007	32.0900	35.1651	39.5496	20.4501	36.6143
2	10.6058	598.9832	83.4273	70.6703	79.0692	-33.6077	41.1263	45.6633	23.6564	15.9619
3	83.4273	83.4273	479.4273	42.1457	87.1770	33.5424	115.6000	98.3505	60.2068	66.7983
4	13.3982	70.6703	42.1457	154.9820	52.5320	10.5183	59.8522	47.8333	21.1194	12.7622
5	30.5007	79.0692	87.1770	52.5320	394.8025	142.9800	130.5297	102.8090	86.1620	412.7222
6	32.0900	-33.6077	33.5424	10.5183	412.9829	820.9829	103.8744	56.4291	38.1224	35.1029
7	35.1651	41.1263	49.8522	59.8522	49.8522	33.8764	49.8522	110.5464	54.7237	54.9689
8	39.5496	45.6633	98.3505	47.8333	60.2068	56.4291	110.5464	116.7097	54.7237	91.1621
9	20.4501	23.6564	60.2068	21.1194	36.6143	38.1224	412.7222	51.7229	43.1334	41.1695

KREUZPRODUKTMATRIX DER GELIEBTEBEN POPULATIONEN

RESM SECTION 1									
ROW	1	2	3	4	5	6	7	8	10
1	455.3853	126.2364	97.3439	15.7236	125.3026	55.4146	91.1563	149.9444	2.5741
2	126.2364	842.7571	43.4287	-14.6892	48.7012	59.7714	41.7867	60.1072	1.6501
3	97.3439	43.4287	391.1907	-6.6120	41.3676	57.3889	-1.3629	60.4503	7.3061
4	15.7236	-14.6892	-6.6120	169.3808	.0521	12.0189	23.2972	21.8054	7.6064
5	125.3026	48.7012	41.3676	.0521	151.6101	59.3806	42.5694	84.1201	15.6995
6	55.4146	59.7714	57.3889	12.0189	59.3806	609.3284	36.2359	81.6702	-12.7575
7	91.1563	41.7867	-1.3629	23.2972	42.5694	36.2359	772.7996	96.6650	20.3602
8	149.9444	60.1072	60.4503	21.8054	84.1201	81.6702	96.6650	233.0283	609.8040
9	7.2840	27.2688	17.4093	14.3829	21.3825	21.5912	-5.4483	609.8040	-9.0908
10	2.5741	1.6501	7.3061	7.6064	15.6995	-12.7575	20.3602	36.4415	155.3639
ALLGEMEINES EIGENWERTPROBLEM Y=V*W+V*L									
SPUR DER MATRIX									
EIGENWERTE									
	6.3002	3.3902	1.5688	1.1744	.8471	.5853	.1841	.0892	.0516

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

37906E+01	.49344F+01	.10191E+02	.43335E+01	.10387E+02	.56214E+01	.11274E+02	.09576E+01	.51052E+01	.76187E+01
14737E+02	.62927E+01	.57365F+01	-1.0851E+01	.78972E+01	.61903E+01	.80687E+01	.11223E+02	.17894E+01	.24564E+01
EIGENVEKTORENMATRIX EV(SPALTEN)									

EV-W SECTION 1

ROW	1	2	3	4	5	6	7	8	10
1	.0029	-.0238	.0066	.0048	.0000	.0102	-.0075	.0006	-.0498
2	-.0003	.0030	-.0218	-.0216	-.0173	-.0010	-.0008	-.0004	-.0010
3	.0005	.0039	-.0089	-.0036	.0178	-.0372	.0186	.0350	-.0020
4	-.0022	.0166	-.0219	-.0260	.0511	.045	-.0018	.0050	-.0003
5	.0059	.0937	.0104	.0176	-.0158	.0101	.0048	.0031	-.0019
6	.0043	-.0042	.0342	-.0275	-.0017	.0006	.0003	.0018	.0011
7	-.0005	.0027	.0008	-.0026	.0073	-.0104	.0070	-.0078	-.0069
8	-.0031	-.0095	-.0138	.0005	.0196	-.0360	-.0098	-.0499	.0274
9	.0018	-.0001	.0002	.0010	.0040	-.0100	-.0373	.0096	.0072
10	.0017	-.0115	-.0023	.0002	-.0014	.0039	.0007	.0016	-.0000

DIVERGENZEN 2	.2331E+01	.1021E+01	.4185E+00	.1727E+01	.7392E+00	.1525E+00	.1832E+01	.2606E+01	.4058E+01	.1109E+02
---------------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

DIVERGENZEN	.1109E+02	.4958E+01	.2606E+01	.2331E+01	.1832E+01	.1525E+00	.1021E+01	.1185E+00	.7392E+01	.1727E+01
-------------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

EIGENVEKTORENMATRIX EV(SPALTEN)

ROW	1	2	3	4	5	6	7	8	9	10
1	-.0490	-.0012	-.0000	-.0029	-.0075	-.0238	.0102	.0066	.0000	.0046
2	-.0010	-.0010	-.0000	-.0003	-.0008	.0030	-.0010	-.0018	-.0173	-.0216
3	-.0056	-.0056	-.0050	-.0005	-.0086	.0039	-.0272	-.0080	.0178	-.0036
4	-.0018	-.0018	-.0050	-.0022	-.0018	.0166	.0445	-.0219	.0511	-.0260
5	-.0019	-.0019	-.0031	-.0059	.0048	.0037	.0101	.0104	-.0158	.0176
6	-.0011	-.0011	-.0018	-.0043	.0003	-.0042	.0006	.0312	-.0017	-.0275
7	-.0049	-.0049	-.0078	-.0005	.0070	-.0027	-.0104	.0008	.0073	-.0046
8	-.0274	-.0339	-.0499	-.0131	-.0095	-.0095	.0360	-.0138	.0196	.0005
9	-.0018	-.0072	-.0096	-.0018	-.0373	-.0001	-.0100	.0002	.0040	.0010
10	-.0000	.0021	.0016	.0817	.0007	-.0115	.0039	-.0023	-.0014	.0002

KREUZPRODUKTMATRIX 5

STGM SECTION 1

ROW	1	2	3	4	5	6	7	8	9	10
1	109.5012	28.3393	6.7435	5.2978	15.8951	6.3791	-.3554	20.0231	-.59.4721	5.4960
2	28.3393	877.0908	20.7604	-26.3761	6.7121	28.7986	-36.9481	37.6605	8.5886	-1.1727
3	6.7435	20.7604	456.5964	.7579	22.3660	40.2317	3.6312	40.0809	13.6733	11.3203
4	5.2978	-26.3761	.7579	125.2500	9.9475	19.8270	51.6945	22.0658	20.7372	8.7564
5	15.8951	6.7121	22.3660	9.9475	55.8983	57.1178	51.5146	94.1748	6.5753	27.6154
6	6.3791	28.7986	40.2317	49.8270	57.1178	170.5482	21.4195	116.5693	24.3123	34.7888
7	-36.9481	-36.9481	-3.6312	51.6945	51.5146	21.4195	910.7791	72.6789	-80.2418	27.1335
8	20.0231	37.6605	40.0809	22.0658	84.1748	116.5693	72.6789	269.0552	-35.4619	53.1160
9	-59.4721	8.7564	13.6733	20.7372	6.5753	24.3123	-80.2418	-35.4619	955.3220	-15.0743
10	5.4960	-1.1727	11.3203	8.7564	27.6154	34.7888	27.1335	53.1160	-15.0743	23.5874

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6	7	8	9	10
1	430.9094	120.3233	406.1644	18.44237	130.1712	63.9850	102.9967	156.4189	33.9244	12.9463
2	120.3233	883.3892	64.3179	17.6573	72.8207	38.9689	67.8451	69.4418	32.3012	7.3618
3	406.1644	64.3179	732.1405	7.1829	62.9716	55.1591	38.2142	79.8736	32.9891	25.7976
4	18.44237	17.6573	7.1829	179.2732	14.2464	8.9167	26.0165	29.3948	14.5036	8.9420
5	130.1712	72.8207	62.9716	179.2732	264.3686	78.0011	65.5081	90.3618	47.9117	44.0677
6	63.9850	38.9689	55.1591	8.9167	78.0011	826.1450	63.7203	61.6233	26.2946	-12.6529
7	102.9967	67.8451	38.2142	26.0165	65.5081	63.7203	63.7203	38.6058	38.6058	28.3370
8	156.4189	69.4418	79.8736	29.3948	90.3618	61.6233	109.2875	178.9203	29.5588	49.1302
9	33.9244	32.3012	32.9891	14.5036	47.9117	26.2946	38.6058	29.5588	339.2097	9.6364
10	12.9463	7.3618	25.7976	9.9420	44.0677	-12.6529	28.3370	49.1302	9.6364	462.5200

ALLGEMEINES EIGENWERTPROBLEM $Y=V \cdot \lambda \cdot V^{-1}$

SPUR DER MATRIX 42.98508

EIGENWERTE

3.6159 3.0463 2.3989 1.6867 1.0335 .6667 .2408 .1750 .0959 .0179

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

EV-M SECTION 1
 1 2 3 4 5 6 7 8 9 10
 1 2331E+01 20836E+01 3887E+01 14655E+01 70241E+01 82672E+01 76778E+01 12628E+02 -12073E+00 5908E+01
 2 1325E+01 7243E+01 3333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 3 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 4 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 5 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 6 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 7 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 8 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 9 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 10 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8	9	10
1	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
2	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
3	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
4	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
5	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
6	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
7	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
8	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
9	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005
10	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005	.0005

EV-M SECTION 1
 1 2 3 4 5 6 7 8 9 10
 1 2331E+01 20836E+01 3887E+01 14655E+01 70241E+01 82672E+01 76778E+01 12628E+02 -12073E+00 5908E+01
 2 1325E+01 7243E+01 3333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 3 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 4 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 5 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 6 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 7 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 8 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 9 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 10 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01

101

EV-M SECTION 1
 1 2 3 4 5 6 7 8 9 10
 1 2331E+01 20836E+01 3887E+01 14655E+01 70241E+01 82672E+01 76778E+01 12628E+02 -12073E+00 5908E+01
 2 1325E+01 7243E+01 3333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 3 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 4 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 5 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 6 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 7 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 8 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 9 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01
 10 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01 1333E+01

EIGENVEKTORENMATRIX EV(CSPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8	9	10
1	.0011	.0033	.0074	.0020	.0005	.0042	.0274	.0235	.0011	.0099
2	.0003	.0010	.0004	.0002	.0025	.0002	.0006	.0076	.0079	.0327
3	.0001	.0001	.0014	.0011	.0098	.0005	.0206	.0041	.0004	.0012
4	.0012	.0021	.0041	.0040	.0048	.0007	.0089	.0020	.0727	.0182
5	.0013	.0043	.0047	.0043	.0025	.0054	.0038	.0010	.0022	.0055
6	.0022	.0035	.0041	.0264	.0023	.0000	.0019	.0004	.0044	.0004
7	.0001	.0022	.0004	.0007	.0082	.0081	.0277	.0341	.0037	.0000
8	.0023	.0063	.0000	.0238	.0489	.0225	.0777	.0311	.0031	.0000
9	.0009	.0002	.0044	.0012	.0060	.0033	.0034	.0123	.0019	.0012
10	.0460	.0033	.0014	.0107	.0001	.0003	.0036	.0023	.0008	.0004

KREUZPRODUKTMATRIX *

SIGM SECTION 1

ROW	1	2	3	4	5	6	7	8	9
1	.2479	.1861	.1733	.0657	.0004	.0066	.0022	.0398	.0239
2	.1861	.9726	.2263	.2263	.0354	.0033	.1521	.1338	.1338
3	.1733	.2263	.2263	.1675	.0160	.0016	.1374	.0622	.0622

1	.0032	.0474	.8332	.1161	.2951	.0023	.1574
2	.0005	.0474	.0161	.17052	.0667	.0108	.0004
3	.0072	.0360	.2511	.0662	.9841	.0609	.0609
4	.0028	.0623	.0023	.0108	.2052	.11967	.0782
5	.0239	.0987	.0576	.0004	.0609	.0783	.12741

KREUZPRODUKTMATRIX DER GLEICHENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6	7	8	9
1	.9233	.0620	.0578	.0219	.0001	.0002	.0024	.0133	.0080
2	.0620	.7078	.0844	.0734	.0118	.0011	.0507	.0713	.0460
3	.0578	.0844	.4100	.0538	.0055	.0339	.0458	.0207	.0329
4	.0219	.0734	.0538	.8246	.0206	.0158	.0120	.0874	.0007
5	.0001	.0118	.0055	.0206	.19165	.0054	.0850	.0008	.0192
6	.0002	.0844	.0339	.0158	.0054	.15805	.0221	.0036	.0001
7	.0507	.0734	.0458	.0120	.0850	.0221	1.4716	.0684	.0201
8	.0713	.0538	.0207	.0874	.0008	.0036	.0684	1.1868	.0261
9	.0080	.0460	.0329	.0007	.0192	.0001	.0203	.0261	.7969

ALLGEMEINES EIGENWERTPROBLEM $Y^*V = \lambda V + L$

SPUR DER MATRIX 12.94436

EIGENWERTE

5.8791 1.8179 1.4753 1.3093 .8246 .7593 .3977 .3303 .1508

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

1	.15886E+00	.26732E+00	.113965E+01	.18458E+00	.10742E-01	.10565E+00	.11910E+00	.15527E+00	.41860E-01
2	.19992E+00	.27634E+00	.27266E+00	.54054E-01	.15799E+00	.54620E+00	.54328E+00	.14900E+00	.4485E-01

EIGENVEKTORENMATRIX FV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8	9
1	.2030	.0509	.1543	.11008	.0368	.0383	.4958	.2619	.8416
2	.4284	.7615	.6851	.1848	.0489	.0632	.2904	.0807	.3582
3	.15604	.2586	.1101	.1656	.0742	.1318	.1340	.0003	.0201
4	.11429	.4622	.3123	.5408	.0547	.4601	.5495	.0124	.3207
5	.0244	.0337	.0076	.0214	.0709	.3839	.2868	.5352	.0137
6	.0587	.0022	.0170	.1023	.7712	.1245	.0889	.0322	.0253
7	.0972	.0701	.1218	.2880	.0929	.5996	.0578	.4308	.1371
8	.1020	.0708	.2399	.7289	.1455	.0070	.4032	.1928	.1814
9	.1703	.7500	.7832	.2029	.0532	.0338	.0412	.0845	.0869

DIVERGENZEN 1 .4306E+01 .1854E+00 .8236E-01 .6124E-01 .2062E-01 .1186E+00 .7937E+00 .9091E+00 .2947E+01

DIVERGENZEN .4306E+01 .2967E+01 .9091E+00 .7937E+00 .1854E+00 .1186E+00 .8236E-01 .6124E-01 .2062E-01

EIGENVEKTORENMATRIX FV(SPALTEN)

EV-M SECTION 1

ALLGEMEINE EIGENWERTPROBLEME
SPUR DER MATRIX
EIGENWERT
EIGENVEKTOR
ANALYSE DER DIVERGENZ
MITTELWERTVEKTOREN DER BEIDEN BLOCKE
EIGENVEKTORPERMUTATION
EV-M SECTION
DIVERGENZEN
DIVERGENZEN
EIGENVEKTORPERMUTATION
EV-M SECTION
KREUZPRODUKTMATRIX
SIGMA SECTION
ROW

59274E-02	10176E+00	11422E+00	25330E+00	75335E+00	59887E+00	2755E-04
2310E+00	7450E+00	2611E+00	2549E+00	3049E+00	1413E+00	
EIGENVEKTORPERMUTATION						
EV-M SECTION						
ROW	1	2	3	4	5	6
1	.0000	-.0000	-.0000	.0000	-.0000	.0000
2	.0000	.0000	-.0000	-.0000	.0000	-.0000
3	.0000	.0000	-.0000	-.0000	.0000	.0000
4	.0000	.0000	-.0000	-.0000	.0000	-.0000
5	-.0000	-.0000	.0000	.0000	-.0000	.0000
6	.0000	-.0000	-.0000	.0000	.0000	.0000
7	.0000	.0000	-.0000	-.0000	.0000	-.0000
8	.0000	.0000	-.0000	-.0000	.0000	.0000
9	.0000	.0000	-.0000	-.0000	.0000	.0000
DIVERGENZEN						
3	1030E+01	7379E+00	6835E+00	2100E+00	2084E+01	1957E+01
3	1030E+01	7379E+00	6835E+00	2100E+00	2084E+01	1957E+01
DIVERGENZEN						
2	2697E+02	4588E+01	3084E+01	1957E+01	7379E+00	5835E+00
2	2697E+02	4588E+01	3084E+01	1957E+01	7379E+00	5835E+00
EIGENVEKTORPERMUTATION						
EV-M SECTION						
ROW	1	2	3	4	5	6
1	.0000	.0000	.0000	.0000	.0000	.0000
2	.0000	-.0000	-.0000	-.0000	.0000	.0000
3	.0000	.0000	.0000	-.0000	.0000	.0000
4	.0000	.0000	.0000	-.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	-.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	-.0000
7	.0000	.0000	.0000	.0000	.0000	.0000
8	.0000	.0000	.0000	.0000	.0000	.0000
9	.0000	.0000	.0000	.0000	.0000	.0000
KREUZPRODUKTMATRIX						
SIGMA SECTION						
ROW	1	2	3	4	5	6
1	1.9446	.0436	-.1697	.2481	-.0065	-.0264
2	.0436	1.3220	.2008	.2787	.0378	-.1218
3	-.1697	.2008	.1700	.1524	-.0550	.0872
4	.2481	.2787	.1524	1.2677	.0168	-.0863
5	-.0065	.0378	-.0550	.0168	.0565	.0252
6	-.0264	-.1218	.0872	-.0863	-.2328	.1365
7	-.0252	.1365	.0459	-.0459	-.0459	.0459
8	.0459	-.0459	.0459	.0459	.0459	.0459
9	.0459	.0459	.0459	.0459	.0459	.0459

105

5	-.0006	.0447	-.0350	-.0279	.4991	.0115	-.1887	-.2181	.0924
6	-.0065	.0078	.0256	.0163	.0115	.3377	.0498	.0049	.0345
7	-.0264	-.1218	.0872	-.0863	-.1887	-.0008	.0000	.2879	-.0658
8	-.0565	-.2328	.0649	-.2465	-.2181	.0049	.2970	.6387	-.1457
9	-.0252	.1365	.0109	.0458	.0921	.0345	-.0658	-.1457	.8972

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6	7	8	9
1	.3578	-.0145	.0566	-.0827	.0002	.0022	.0088	-.0188	.0084
2	-.0145	.5913	-.0669	-.0929	-.0149	-.0026	.0406	.0776	-.0455
3	.0566	-.0669	1.0236	-.0508	.0184	-.0085	-.0291	-.0216	-.0036
4	-.0827	-.0929	.0508	.6358	.0326	-.0056	.0288	.0822	-.0153
5	.0002	-.0149	.0184	.6358	2.0379	.0336	.0629	.0727	.0307
6	.0022	-.0026	-.0085	-.0056	-.0038	1.9030	-.0066	-.0016	.0115
7	.0088	.0406	-.0291	.0288	.0727	-.0066	1.6663	.0960	.0219
8	-.0188	.0776	-.0216	.0822	.0727	-.0016	-.0960	1.3493	.0486
9	.0084	-.0455	-.0036	-.0153	-.0387	-.0115	.0219	.0486	.9232

ALLGEMEINES EIGENWERTPROBLEM Y+V=W+V*L

SPUR DER MATRIX 13.21435

EIGENWERTE

6.2961	3.3824	1.4120	.9945	.6196	.1841	.1552	.1185	.0519
--------	--------	--------	-------	-------	-------	-------	-------	-------

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

.13175E+00	.37125E+00	-.14723E+00	.33633E+00	.25835E+00	.52777E-01	-.49852E+00	-.67308E+00	.26786E+00	-.52004E-01
.16327E+00	-.69272E+00	.22208E+00	-.12992E+00	.17562E+00	-.34033E+00	-.27716E+00	.45768E-01		
EIGENVEKTORENMATRIX FV(SPALTEN)									

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8	9
1	-1.6247	.4690	.1962	-.0170	-.0111	-.0156	.0125	-.0477	-.0993
2	-.3032	-1.0411	.6849	.3319	-.0594	.0262	.0069	.0432	.1521
3	-.0673	-.2403	-.0446	.1483	-.2885	-.1758	-.1698	-.2191	-.8484
4	-.5619	-.7079	-.9132	-.0122	-.0802	-.0170	-.1075	.0491	.1476
5	.0040	-.0210	.0847	-.1394	.2716	-.2984	.5128	-.1703	.1048
6	.0003	-.0118	.0033	-.0193	-.0438	.4358	.3124	.1397	.0422
7	.0462	.1037	-.0273	.1430	-.3700	-.0373	.0824	-.5985	.2603
8	.0542	.2680	.0234	.2616	-.5519	-.0948	-.3749	.3853	.1455
9	-.0227	-.2266	.1907	-.9074	-.4218	.0510	-.0092	.0272	.0077

DIVERGENZEN 2	.2330F+01	.1015E+01	.7275E-01	.5496E-01	.1335E+00	.1832E+01	.2456E+01	.3572E+01	.1113E+02
---------------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

DIVERGENZEN	.1113E+02	.3572E+01	.2456E+01	.2330E+01	.1832E+01	.1015E+01	.1335E+00	.7275E-01	.5496E-01
-------------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

EIGENVEKTORENMATRIX FV(SPALTEN)

EV-M SECTION 1

6
7
8

-.0666
.3148
.2291
.0796
-.1091

-.2502
-.0422
-.1472
-.3006
-.5252
-.1426
-.0076
-.0672
-.0779

DIVERGENZEN
6.4235E+04 2.203E+00 1.5399E+00 1.7428E+00 1.2741E+00

DIVERGENZEN
6.4235E+04 2.203E+00 1.5399E+00 1.7428E+00 1.2741E+00

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8
1	1.5945	-.0496	-.1082	-.0911	.3235	.0098	.0752	-.1257
2	.3499	.0214	-.0295	.1642	-.8466	.2396	.6444	-.0517
3	.2296	.1078	.5434	-.5348	-.4345	-.0860	-.5227	-.3055
4	.0078	-.5621	.1092	.0495	-.0499	-.1043	-.0231	.0168
5	.1791	.0173	.1315	-.3543	-.0610	.0482	.0562	1.0096
6	-.0666	-.1420	-.5252	-.3006	-.1472	.3148	.2502	-.0422
7	.3113	.0672	.0076	.6313	-.2234	.2291	-.6915	.2318
8	.1008	.0779	-.3369	.0208	-.2440	-.8796	-.1081	.0739

KREUZPRODUKTMATRIX 2

SIGM SECTION 1

ROW	1	2	3	4	5	6	7	8
1	.0519	-.0090	.0000	-.0000	.0000	.0000	-.0000	.0000
2	-.0000	.1185	.0000	.0000	.0000	.0000	.0000	.0000
3	-.0000	.0000	.1552	.0000	.0000	.0000	.0000	.0000
4	-.0000	.0000	.0000	.6.2960	.0000	.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	.1841	.0000	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	3.3824	.0000	.0000
7	-.0000	.0000	.0000	.0000	.0000	.0000	.6196	.0000
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	1.4120

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6	7	8
1	1.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2	.0000	1.0000	.0000	.0000	.0000	.0000	.0000	.0000
3	.0000	.0000	1.0000	.0000	.0000	.0000	.0000	.0000
4	.0000	.0000	.0000	1.0000	.0000	.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	1.0000	.0000	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	1.0000	.0000	.0000
7	.0000	.0000	.0000	.0000	.0000	.0000	1.0000	.0000
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	1.0000

ALLGEMEINES EIGENWERTPROBLEM Y*V=W*Y*L

SPUR DER MATRIX 12.21986

EIGENWERTE 6.2961 3.3824 1.4120 .6196 .1841 .1552 .1185 .0519

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOR DER GELIEBEN POPULATIONEN

.21825E+04 .08415E+01 .85358E+04 .10601E+02 .32539E+00 .51395E+01 .5157E+00 .31702E+00
 .29413E+03 .16100E+00 .90493E+04 .39071E+03 .70293E+01

EIGENVEKTORMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8
1	-.0000	.0000	-.0000	-.0000	.0000	.0000	-.0000	1.0000
2	.0000	.0000	-.0000	.0000	.0000	.0000	1.0000	.0000
3	.0000	.0000	.0000	.0000	.0000	-.1.0000	-.0000	.0000
4	-1.0000	.0000	-.0000	.0000	.0000	.0000	-.0000	.0000
5	.0000	.0000	.0000	.0000	-.1.0000	.0000	-.0000	.0000
6	.0000	-1.0000	.0000	.0000	.0000	.0000	-.0000	.0000
7	.0000	.0000	.0000	1.0000	.0000	.0000	.0000	.0000
8	.0000	.0000	-1.0000	-.0000	-.0000	-.0000	.0000	.0000

DIVERGENZEN 2 .2330E+01 .1015E+01 .7275E+01 .1335E+01 .1832E+01 .2456E+01 .3572E+01 .1113E+02

DIVERGENZEN 1 .1413E+02 .3572E+01 .2456E+01 .2330E+01 .1832E+01 .1015E+01 .1335E+00 .7275E+01

EIGENVEKTORMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8
1	1.0000	-.0000	.0000	-.0000	.0000	.0000	-.0000	-.0000
2	.0000	1.0000	-.0000	.0000	.0000	.0000	.0000	.0000
3	-.0000	.0000	-1.0000	.0000	.0000	.0000	.0000	.0000
4	.0000	.0000	.0000	-1.0000	.0000	.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	-1.0000	.0000	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	-1.0000	.0000	.0000
7	.0000	.0000	.0000	.0000	.0000	.0000	1.0000	.0000
8	-.0000	.0000	.0000	.0000	-.0000	.0000	-.0000	-1.0000

KREUZPRODUKTMATRIX 3

SIGM SECTION 1

ROW	1	2	3	4	5	6	7	8
1	.6229	-.2788	-.1627	.0438	-.2182	.1334	-.2153	.0170
2	-.2788	1.2868	.0944	-.0422	-.0367	.0338	-.0065	.0003
3	-.1627	.0944	1.5097	-.0250	-.0155	-.0979	-.2532	-.1450
4	.0438	-.0422	-.0250	.4229	.0194	.1213	-.0759	.0609
5	-.2103	.0367	-.0155	-.0194	1.5776	.0040	-.0988	.0730
6	.1334	.0338	-.0979	.1213	.0040	.2626	-.2626	.0172
7	-.2153	-.0065	-.2532	-.0759	-.0988	-.2626	1.2547	.0472
8	.0170	-.0003	-.1650	.0609	-.0730	.0172	.0472	.2464

KREUZPRODUKTMATRIX DER GELIEBENEN POPULATIONEN

ROW	1	2	3	4	5	6	7	8
1	.876	.0280	.0342	.0000	.0701	-.0445	.0718	-.0057
2	.0930	.0314	-.0314	.0141	.0123	-.0113	.0022	.0001
3	.0540	.0545	.0083	.0083	.0052	.0326	.0844	.0550
4	-.0046	.0141	.0083	.0083	.0065	-.0013	.0253	-.0203
5	.0701	.0123	.0052	.0065	.0355	-.0013	.0329	.0243
6	-.0445	-.0113	.0326	-.0065	-.0013	1.9968	.0875	-.0057
7	.0718	.0022	.0844	.0253	.0329	.0875	.7883	.0157
8	-.0057	.0001	.0550	-.0023	.0243	-.0057	-.0157	1.3886

ALLGEMEINES EIGENWERTPROBLEM $Y \cdot V = \lambda \cdot V + L$

SPUR DER MATRIX 10.98764

EIGENWERTF 3.4657 3.0443 2.3318 1.5627 .2516 .1751 .1386 .0180

ANALYSE DER DIVERGENZEN

MITTELMERTVEKTOREN DER BEIDEN POPULATIONEN

1	.15739E+00	.36439E+00	.43233E+00	-.23293E+00	-.24413E-01	-.46868E+00	.49256E+00	-.11987E+00	.57541E+00	.21829E+00
2	.17964E+00	-.28773E+00	.99177E-01	-.42265E+00	.39291E+00	-.13210E-01				

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8
1	-.2764	-.1043	.4089	.2612	-.8610	-.1304	-.3131	.0878
2	.4246	-.2422	-1.0712	.4776	-.2148	-.0438	-.0862	.0460
3	1.1116	-.5545	.2840	-.4627	-.1661	.1119	-.0365	.0114
4	-.0077	-.0013	.0221	.0282	.0107	.1879	.1028	.5291
5	.5827	1.2327	-.0124	-.0371	-.1623	.0222	-.0674	.0152
6	-.0075	.0064	.0533	.1831	-.2159	.3333	.5150	.2124
7	-.4162	.0723	.5149	-.8694	-.3121	.0682	.0744	.0090
8	-.1185	.0016	-.0360	.0372	.0723	.6770	-.4728	-.1428

DIVERGENZEN 3 .1007E+01 .7269E+00 .5127E+00 .1924E+00 .1862E+01 .1949E+01 .2945E+01 .2685E+02

DIVERGENZEN 2 .2685E+02 .2945E+01 .1949E+01 .1862E+01 .1007E+01 .7269E+00 .5127E+00 .1924E+00

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8
1	.0978	-.3131	-.1304	-.8610	-.2764	-.1043	.4089	.2612
2	.0460	-.0862	-.0438	-.4246	.4246	-.2422	-.0862	.4627
3	-.0114	-.0365	.1119	-.1661	1.1116	-.5545	.2840	-.4627
4	.5291	.1028	.1879	.0107	-.0077	.0013	.0221	.0282
5	.0152	-.0674	.0227	-.1623	.5827	1.2327	-.0124	-.0371
6	-.2124	.5150	.3333	-.2159	-.0075	.0064	.5150	.1831
7	.0090	.0744	.0682	-.3121	-.4162	.0723	.0533	-.8694
8	-.1428	-.4728	.6770	.0723	-.1185	.0016	-.0360	.0372

SIGN SECTION 1

ROW	1	2	3	4	5	6	7	8
1	.0519	-.0000	.0000	-.0000	.0000	.0000	-.0000	.0000
2	-.0000	.1185	.0000	.0000	.0000	.0000	.0000	.0000
3	.0000	.0000	.1552	.0000	.0000	.0000	.0000	.0000
4	-.0000	.0000	.0000	.62960	.0000	.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	.1841	.0000	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	.37824	.0000	.0000
7	-.0000	.0000	.0000	.0000	.0000	.0000	.6196	.0000
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	1.4120

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6	7	8
1	1.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2	.0000	1.0000	.0000	.0000	.0000	.0000	.0000	.0000
3	.0000	.0000	1.0000	.0000	.0000	.0000	.0000	.0000
4	.0000	.0000	.0000	1.0000	.0000	.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	1.0000	.0000	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	1.0000	.0000	.0000
7	.0000	.0000	.0000	.0000	.0000	.0000	1.0000	.0000
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	1.0000

ALLGEMEINES EIGENWERTPROBLEM Y=V*V*L

SPUR DER MATRIX 12.21986

EIGENWERTE 6.2961 3.5824 1.4120 .6196 .1552 .1185 .0519

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

.21625E-01	.68215E-01	.85833E-01	-.59005E+00	.30610E-02	-.82529E+00	.50280E+00	.51395E-01	.31702E+00
.29113E+00	-.16869E+00	.90019E-01	-.30378E+00	.38949E+00	-.70298E-01			

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8
1	-.0000	.0000	-.0000	-.0000	.0000	.0000	-.0000	1.0000
2	.0000	.0000	.0000	.0000	.0000	.0000	1.0000	.0000
3	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
4	-.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

DIVERGENZEN 2 .2330E+01 .1015E+01 .7274E-01 .1335E+00 .1832E+01 .2456E+01 .3572E+01 .1113E+02

DIVERGENZEN .1113E+02 .3572E+01 .2456E+01 .2330E+01 .1832E+01 .1015E+01 .1335E+00 .7275E-01

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7	8
1	1.0000	-.0000	.0000	-.0000	.0000	.0000	-.0000	.0000
2	-.0000	1.0000	-.0000	.0000	.0000	.0000	.0000	-.0000
3	-.0000	-.0000	1.0000	-.0000	.0000	.0000	.0000	.0000
4	-.0000	-.0000	-.0000	1.0000	.0000	.0000	.0000	.0000
5	-.0000	-.0000	-.0000	-.0000	1.0000	.0000	.0000	.0000
6	-.0000	-.0000	-.0000	-.0000	-.0000	1.0000	.0000	.0000
7	-.0000	-.0000	-.0000	-.0000	-.0000	-.0000	1.0000	.0000
8	-.0000	.0000	-.0000	.0000	-.0000	.0000	-.0000	-1.0000

KREUZPRODUKTMATRIX 1

SIGM SECTION 1

ROW	1	2	3	4	5	6	7
1	1.8424	.2899	-.1808	.0114	-.1377	.1605	.2845
2	.2899	.8088	.0249	.0143	-.0562	.0434	-.0708
3	-.1808	.0269	.6025	.0779	-.0578	.1775	-.2136
4	.0114	.0143	.0779	.5588	.0128	-.1781	.0327
5	-.1377	-.0562	-.0578	.0128	.6691	.0219	-.0896
6	.1605	.0434	.1275	-.1781	.0219	1.0885	.2324
7	.2845	-.0708	-.2136	-.0327	-.0896	.2324	.8683

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6	7
1	.4033	-.0966	.0602	-.0038	.0459	-.0535	-.0948
2	-.0966	.7699	-.0089	-.0048	.0187	-.0145	.0236
3	.0602	-.0089	.8509	-.0260	.0192	-.0425	.0712
4	-.0038	-.0048	-.0260	2.9124	-.0043	.0594	.0109
5	.0459	.0187	.0192	-.0043	.8383	.0073	.0299
6	-.0535	-.0145	-.0425	.0594	.0073	1.7646	.0775
7	-.0948	.0236	.0712	.0109	.0299	.0775	.9171

ALLGEMEINES EIGENWERTPROBLEM Y=V=**V*L
SPUR DER MATRIX 9.95250
EIGENWERTE 5.7885 1.3738 .8548 .8156 .5120 .4463 .1615

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

.18294E+01	.33104E+00	-.18630E+00	.56393E-02	-.15646E+00	.86653E-01	.26551E+00	.79837E-01	.22939E+00	-.25766E+00
.36350E+00	-.38879E-01	.55000E+00	.48859E+00						

EIGENVEKTORENMATRIX FV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7
1	-1.5933	.0772	.2799	.1095	-.1868	.1050	-.0506
2	-.3582	.5996	-.6270	-.6479	.1166	-.1658	.0088
3	.7227	.5056	-.0726	.4201	-.6884	-.4683	-.1451
4	.0030	.0682	-.0219	-.0167	-.1764	.0642	.5544

5	.1795	-.0459	.0077	-.2299	-.3384	-.0244
6	-.0558	.2941	.2340	.4700	-.3527	.1821
7	-.3158	-.7184	-.3625	-.0013	-.6303	.0619

DIVERGENZEN 1 .4495E+01 .8740E-01 .4244E-01 .2931E-01 .4249E+00 .5890E+00 .2657E+01

DIVERGENZEN 2 .6096E+01 .2657E+01 .5890E+00 .4649E+00 .8740E-01 .4244E-01 .2931E-01

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7
1	-.15933	-.0506	.1050	-.1868	.0732	.2799	.1095
2	-.3582	.0088	-.1658	.1166	.5996	-.6270	-.6479
3	.2227	-.1431	-.4683	-.6884	.5056	-.0726	.4201
4	.0030	.5546	.0642	-.1764	.0082	-.0219	-.0167
5	.1795	-.0246	-.3384	-.2399	-.0469	.7077	-.7083
6	-.0558	.1821	-.3527	.4700	.2941	.7340	.2215
7	-.3158	.0619	-.6303	-.0082	-.7184	-.3425	.0018

KREUZPRODUKTMATRIX 2

SIGM SECTION 1

ROW	1	2	3	4	5	6	7
1	.0519	-.0000	.0000	-.0000	-.0000	.0000	.0000
2	-.0000	.1185	.0000	.0000	.0000	-.0000	.0000
3	.0000	.0000	.1552	.0000	.0000	.0000	.0000
4	-.0000	.0000	.0000	.62961	.0000	.0000	.0000
5	.0000	.0000	.0000	-.0000	.1841	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	3.3824	.0000
7	.0000	.0000	.0000	.0000	.0000	.0000	.6196

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6	7
1	1.0000	-.0000	-.0000	-.0000	-.0000	-.0000	-.0000
2	-.0000	1.0000	.0000	.0000	.0000	.0000	.0000
3	.0000	.0000	1.0000	.0000	.0000	.0000	.0000
4	.0000	.0000	.0000	1.0000	.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	1.0000	.0000	.0000
6	.0000	.0000	.0000	.0000	.0000	1.0000	.0000
7	.0000	.0000	.0000	.0000	.0000	.0000	1.0000

ALLGEMEINES EIGENWERTPROBLEM YAV=WW*Y+L

SPUR DER MATRIX 40.80787

EIGENWERTE 6.2961 3.3824 .6196 .1841 .1552 .1185 .0519

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

-0.15738E+00 .36438E+00 -0.43234E+00 .23293E+00 .21827E+00 -0.17564E+00
 .28773E+00 -0.99470E-01 .42266E+00 .39290E+00
 EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7
1	-0.2816	.1036	.4055	.2620	-0.8584	.3415	-0.1086
2	.4423	.2435	-1.0629	.4795	-0.2134	.0975	-0.0483
3	-1.1029	-0.5568	-0.2947	.4728	.1763	.0305	.0040
4	.0064	-0.0013	-0.0227	-0.0274	-0.0025	.2041	.5332
5	-0.5851	1.2314	.0044	.0412	.1660	-0.0449	.0235
6	.0067	-0.0064	-0.0540	-0.1828	.2234	.6049	-0.2886
7	-0.4142	-0.0732	-0.5249	-0.8640	-0.3135	-0.0956	-0.0061

DIVERGENZEN 3 .9734E+00 .7272E+00 .5101E+00 .1905E+00 .1878E+01 .2619E+01 .2167E+02

DIVERGENZEN .2167E+02 .2619E+01 .1878E+01 .9734E+00 .7272E+00 .5101E+00 .1905E+00

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6	7
1	-0.1086	.3415	-0.8584	-0.2816	.1036	.4055	.2620
2	-0.0483	.0975	-0.2134	.4728	-0.0629	-0.0025	.2620
3	.0040	.0305	.1763	-1.1029	-0.5568	.2987	.4728
4	.5332	.2041	-0.0025	.0064	-0.0013	-0.0227	-0.0274
5	.0235	-0.0449	.1660	-0.5851	1.2314	.0044	.0412
6	-0.2886	.6049	.2234	-0.0067	.0064	-0.0540	.1828
7	-0.0061	-0.0956	-0.3135	-0.4142	-0.0732	-0.5249	-0.8640

KREUZPRODUKTMATRIX 1

SIGN SECTION 1

ROW	1	2	3	4	5	6	7
1	1.8424	.2899	-0.1808	.0114	-0.1377	.1605	.2845
2	.2899	.8088	.0249	.0143	-0.0562	.0434	-0.0708
3	-0.1808	.0249	.6025	.0779	-0.0578	.1275	-0.2136
4	.0114	.0143	.0779	.5588	.0128	-0.1781	-0.0327

5
6
7

-.0210
1.0885
-.2324

.0091
-.0219
-.0896

-.1781
-.0327

.0073
-.1275
-.2136

-.0434
-.0708

-.1377
-.1605
-.2845

PRODUKT-MATRIX DER SPATIELENE DER LATIONEN

RESM SECTION 1

ROW 1 2 3 4 5 6 7

-.0966
-.0966
-.0602
-.0038
-.0535
-.0145
-.0236

-.0089
-.0089
-.0509
-.0260
-.0425
-.0192
-.0775

-.0048
-.0048
-.0124
-.0043
-.0073
-.0073
-.0171

-.0145
-.0145
-.0192
-.0594
-.0299
-.0775
-.0171

-.0966
-.0966
-.0602
-.0038
-.0535
-.0145
-.0236

ALLGEMEINES EIGENWERTPROBLEM $Y' = AY + V + L$

SPUR DER MATRIX 9.93250

EIGENWERTE 5.7885 1.3738 .8548 .8156 .5120 .6463 .1615

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

116

-.22939E+00

.79837E-01

.26551E+00

.86653E-01

-.15646E+00

.56393E-02

-.18630E+00

-.58000E+00

.46859E+00

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW 1 2 3 4 5 6 7

-.15933
-.3582
-.2227
-.0030
-.1795
-.0558
-.3158

.0772
.5996
.5056
.0082
-.0469
.2941
-.7184

.2799
-.6270
-.0726
-.0219
-.7077
-.2340
-.3425

.1095
-.6479
-.4201
-.0167
-.7083
-.2215
-.0018

-.1808
-.1166
-.6884
-.1764
-.2299
-.4700
-.0082

.1050
-.1658
-.4683
-.0642
-.3384
-.3527
-.6303

-.0506
-.0088
-.1451
-.5546
-.0246
-.1821
-.0619

DIVERGENZEN 1 .4196E+01 .8740E-01 .4244E-01 .2931E-01 .4648E+00 .5890E+00 .2657E+01

DIVERGENZEN .4196E+01 .2657E+01 .5890E+00 .4648E+00 .8740E-01 .4244E-01 .2931E-01

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW 1 2 3 4 5 6 7

-.15933
-.3582
-.2227
-.0030
-.1795
-.0558
-.3158

.0772
.5996
.5056
.0082
-.0469
.2941
-.7184

.2799
-.6270
-.0726
-.0219
-.7077
-.2340
-.3425

.1095
-.6479
-.4201
-.0167
-.7083
-.2215
-.0018

-.1808
-.1166
-.6884
-.1764
-.2299
-.4700
-.0082

.1050
-.1658
-.4683
-.0642
-.3384
-.3527
-.6303

-.0506
-.0088
-.1451
-.5546
-.0246
-.1821
-.0619

KREZPRODUKTMATRIX :

SIGH SECTION 4

ROW	1	2	3	4	5	6
1	5.7885	-.0000	.0000	-.0000	.0000	.0000
2	-.0000	.1615	.0000	-.0000	.0000	-.0000
3	.0000	.0000	.4848	-.0000	.0000	-.0000
4	-.0000	-.0000	-.0000	.5120	.0000	.0000
5	.0000	.0000	.0000	.0000	1.3738	.0000
6	.0000	-.0000	-.0000	.0000	.0000	.8548

KREZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 4

ROW	1	2	3	4	5	6
1	1.0000	-.0000	.0000	-.0000	.0000	.0000
2	-.0000	1.0000	-.0000	.0000	-.0000	.0000
3	.0000	-.0000	1.0000	-.0000	.0000	-.0000
4	-.0000	.0000	-.0000	1.0000	-.0000	.0000
5	.0000	-.0000	.0000	-.0000	1.0000	.0000
6	.0000	.0000	-.0000	.0000	.0000	1.0000

ALLGEMEINES EIGENWERTPROBLEM Y=VW+V+L

SPUR DER MATRIX

EIGENWERTE 5.7886 1.3738 .8548 .5120 .4463 .1615

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

-.23950E+01 .18642E+02 .27373E-01 -.79240E-02 .49016E-01 -.34404E-02 -.45131E+00 .36710E+00 -.36181E+00 .38869E+00

-.18607E+00 -.17003E+00

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 4

ROW	1	2	3	4	5	6
1	-1.0000	.0000	.0000	.0000	-.0000	-.0000
2	-.0000	.0000	.0000	-.0000	.0000	-1.0000
3	-.0000	-.0000	.0000	.0000	1.0000	.0000
4	-.0001	.0000	-.0000	1.0000	.0000	.0000
5	-.0000	1.0000	.0000	-.0000	.0000	.0000
6	-.0001	-.0000	-1.0000	-.0000	.0000	-.0000

DIVERGENZEN 1.170E+01 .8740E-01 .4244E-01 .4548E+00 .5890E+00 .2656E+01

DIVERGENZEN .4996E+01 .2654E+01 .5890E+01 .4548E+00 .8740E-01 .4244E-01

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6
1	-1.0000	-.0000	-.0000	-.0000	.0000	.0000
2	-.0000	-1.0000	-.0000	-.0000	.0000	.0000
3	-.0000	-.0000	1.0000	-.0000	-.0000	.0000
4	-.0000	-.0000	-.0000	1.0000	.0000	-.0000
5	-.0000	-.0000	-.0000	-.0000	1.0000	.0000
6	-.0000	-.0000	-.0000	-.0000	-.0000	-1.0000

KREUZPRODUKTMATRIX 2

SIGM SECTION 1

ROW	1	2	3	4	5	6
1	.2331	-.0391	.1621	-.1114	.0693	.0467
2	-.0389	2.0549	-.0056	-.3097	.1715	.0518
3	.1621	-.0056	.7519	-.5679	-.1122	-.1792
4	-.1114	-.3097	-.5679	1.0300	.4175	.3644
5	.0693	.1715	-.1122	.4175	.6957	.3288
6	.0467	.0518	-.1792	.3644	.3288	.4045

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6
1	2.8517	.0127	-.0540	.0371	-.0231	-.0156
2	.0127	.3689	.0019	.1032	-.0572	-.0172
3	-.0540	.0019	.8981	.1893	.0374	.0597
4	.0371	.1032	.1893	.8273	-.1392	-.1215
5	-.0231	-.0572	.0374	-.1392	1.2260	1.1096
6	-.0156	-.0172	.0597	-.1215	1.1096	1.1501

ALLGEMEINES EIGENWERTPROBLEM Y*V=W*V*Lambda

SPUR DER MATRIX 10.14882

EIGENWERTE 6.2864 2.8871 .6196 .1650 .1379 .0529

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

-.2846E+00	.52073E+00	-.53790E+00	.34341E+00	-.11438E+00	-.24635E-01	-.11557E+01	.19414E+00	-.17339E+00	.27157E+00
-.10220E+00	-.16297E+00								

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6
1	.0022	-.0411	-.1301	-.1212	.0592	-.5609
2	1.6386	.3341	.0710	.1251	.0625	-.0441
3	.1580	-.6587	-.6624	.4752	.2447	.0809
4	-.4468	.8638	-.0080	.6144	.1830	-.1411
5	.0276	.4273	-.5943	-.1439	-.5360	.0751
6	-.0441	.3962	-.3212	-.4532	.6268	.2065

DIVERGENZEN 2 ,2335E+01 ,7383E+00 ,1333E+00 ,2342E+00 ,2720E+00 ,44E+02

DIVERGENZEN ,1444E+02 ,2720E+00 ,2734E+01 ,2335E+00 ,2720E+00 ,44E+02

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6
1	-.5609	.0592	-.1212	.0022	-.0411	-.1301
2	-.0441	.0625	.1251	1.6386	.3341	.0710
3	.0809	.2447	.4752	.1580	-.6587	-.6624
4	-.1411	.1830	.6144	-.4468	.8638	-.0080
5	.0751	-.5360	-.1439	.0276	.4273	-.5943
6	.2065	.6268	-.4532	-.0441	.3962	-.3212

KREUZPRODUKTMATRIX 3

SIGM SECTION 1

ROW	1	2	3	4	5	6
1	.6672	.0194	-.0646	.1285	-.0068	-.1543
2	.0194	.1381	-.0313	.2678	-.1430	-.0890
3	-.0646	-.0313	1.3625	.5337	.1989	.1070
4	.1285	.2678	.5337	1.0544	-.3182	-.3530
5	-.0068	-.1430	.1979	-.3182	1.1568	-.3275
6	-.1543	-.0890	.1070	-.3530	-.3275	1.6006

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6
1	2.7070	-.0065	.0215	-.0428	.0016	.0515
2	-.0065	1.0078	.0104	-.0893	.0477	.0297
3	.0215	.0104	.6946	-.1779	-.0660	-.0357
4	-.0428	-.0893	-.1779	.8192	.1061	.1177
5	.0016	.0477	-.0660	.1061	1.0723	.1092
6	.0515	.0297	-.0357	.1177	.1092	.7514

ALLGEMEINES EIGENWERTPROBLEM Y*V=W*V*L

SPUR DER MATRIX 7.96369
EIGENWERTE 3.3774 2.4486 1.5843 .2974 .2234 .0329

ANALYSE DER DIVERGENZEN

MITTELWERTVEKTOREN DER BEIDEN POPULATIONEN

-.15264E+00 .31833E+00 -.34353E+00 .53904E+00 -.22754E+00 -.28799E+00 -.11987E+01 .26161E+00 -.23818E+00 .20637E+00

-.64484E+01 -.75185E+01

EIGENVEKTORENMATRIX EV(SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5
1	-.0357	.0523	-.141	-.2541	.548
2	-.1588	.0273	-.2707	-.2751	-.2473
3	-.8593	-.6192	-.961	.3242	.8601
4	-.8511	-.1180	-.4756	-.4552	-.2161
5	-.0501	.2861	.7873	-.4706	-.1839
6	.5548	-.9788	.0244	-.3302	-.0492

DIVERGENZEN 3 .9155E+00 .4707E+00 .2418E+00 .1034E+01 .2086E+01 .1423E+02

DIVERGENZEN .1423E+02 .2065E+01 .1054E+01 .0155E+00 .4707E+00 .2418E+00

EIGENVEKTORENMATRIX EV (SPALTEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6
1	-.0580	.5444	-.2541	-.0357	.0533	-.0446
2	-.8782	-.2473	-.2751	-.1588	.0273	-.2707
3	-.1756	.1660	.3294	-.8593	-.6192	.4901
4	.3565	-.2161	-.4552	-.8511	-.1180	-.4756
5	.0311	-.1839	-.4706	-.0501	.2861	.7873
6	.0439	-.0692	-.3302	.5548	-.9788	.0244

KREUZPRODUKTMATRIX 1

SIGM SECTION 1

ROW	1	2	3	4	5	6
1	5.7885	-.0000	.0000	-.0000	.0000	.0000
2	-.0000	.1615	.0000	-.0000	.0000	-.0000
3	.0000	.0000	.4443	-.0000	.0000	-.0000
4	-.0000	.0000	-.0000	.5120	.0000	.0000
5	.0000	.0000	.0000	.0000	1.3738	.0000
6	.0000	-.0000	-.0000	.0000	.0000	.8548

KREUZPRODUKTMATRIX DER GEBLIEBENEN POPULATIONEN

RESM SECTION 1

ROW	1	2	3	4	5	6
1	1.0000	-.0000	.0000	-.0000	.0000	.0000
2	-.0000	1.0000	.0000	.0000	.0000	.0000
3	.0000	.0000	1.0000	.0000	.0000	.0000
4	-.0000	.0000	.0000	1.0000	.0000	.0000
5	.0000	.0000	.0000	.0000	1.0000	.0000
6	.0000	.0000	.0000	-.0000	.0000	1.0000

ALLGEMEINES EIGENWERTPROBLEM Y*V=W*V*L
SPUR DER MATRIX 9.13701
EIGENWERTE 5.7886 1.3738 .8548 .5120 .4463 .1615

MITTELWERTVEKTOREN DES RITTER MOD. 1.3.1964

23950E+04 18662E+02 2371E+01 340E+02 45131E+00 36710E+00 -36481E+00 3886YE+00

EIGENVEKTORENMATRIX EV(SPALEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6
1	-1.0000	.0000	.0000	.0000	-.0000	-.0000
2	-.0000	-.0000	.0000	-.0000	.0000	-1.0000
3	-.0000	-.0000	.0000	-.0000	1.0000	.0000
4	-.0001	.0000	-.0000	1.0000	.0000	-.0000
5	-.0000	1.0000	.0000	-.0000	.0000	-.0000
6	-.0001	-.0000	-1.0000	-.0000	.0000	-.0000

DIVERGENZEN 1 .4196E+01 .8740E-01 .4244E-01 .4648E+00 .5890E+00 .2656E+01

DIVERGENZEN .4196E+01 .2656E+01 .5890E+00 .4648E+00 .8740E-01 .4244E-01

EIGENVEKTORENMATRIX EV(SPALEN)

EV-M SECTION 1

ROW	1	2	3	4	5	6
1	-1.0000	-.0000	-.0000	.0000	.0000	.0000
2	-.0000	-1.0000	.0000	-.0000	-.0000	.0000
3	-.0000	.0000	1.0000	-.0000	-.0000	.0000
4	-.0001	-.0000	.0000	1.0000	.0000	-.0000
5	-.0000	.0000	.0000	-.0000	1.0000	.0000
6	-.0001	-.0000	.0000	-.0000	-.0000	-1.0000

DISKRIMINANZANALYSE

ANZAHL DER GRUPPEN 3
ANZAHL DER VARIABLEN 5
ANZAHL DER OBJEKTE 600

KREUZPRODUKTMATRIX W (INNERHALB DER GRUPPEN)

W-M SECTION 1					
ROW	1	2	3	4	5
1	127.5031	27.1640	6.1102	-26.6051	-17.2524
2	27.1640	297.3042	-52.8847	58.8557	-24.0645
3	6.1102	-52.8847	722.8857	50.3830	-8.2888
4	-26.6051	58.8557	50.3830	328.1814	39.2543
5	-17.2524	-24.0645	-8.2888	39.2513	473.8594

KREUZPRODUKTMATRIX B (ZWISCHEN DEN GRUPPEN)

B-M SECTION 1					
ROW	1	2	3	4	5
1	475.5762	89.0090	100.0646	-99.7136	49.0026
2	89.0090	20.5186	22.5019	-15.7905	7.4698
3	100.0646	22.5019	24.7441	-18.1724	8.6469
4	-99.7136	-15.7905	-18.1724	23.0442	-11.5406
5	49.0026	7.4698	8.6469	-11.5406	5.7993

ALLGEMEINES EIGENWERTPROBLEM B*T = W*T*L

SPUR DER MATRIX 3.86437
EIGENWERTE 3.8327 .0321 .0000 -.0000 -.0000

WILKS LAMBDA KRITERIUM
TEST AUF VERSCHIEDENHEIT DER GRUPPEN

WILKS LAMBDA = .2005008

FREIHEITSGRADE

F1 = 10.00 F2 = 886.00

F-WERT = 109.2678

EIGENVEKTORENMATRIX T (ZEILEN)

T-M SECTION 1					
ROW	1	2	3	4	5
1	.0873	.0010	.0062	-.0019	.0059
2	.0123	-.0414	-.0379	-.0119	.0090

MATRIX DER GEWICHTETEN EIGENVEKTOREN G (ZEILEN)

G-Matrix SECTION 1

ROW	1	2	3	4	5
1	.9846	.0174	.1118	-.0340	.1286
2	.1341	-.6880	-.6553	-.2082	.1878

MATRIX DER ORTHOGONALISIERTEN EIGENVEKTOREN V (ZEILEN)

V-Matrix SECTION 1

ROW	1	2	3	4	5
1	.9949	.0115	.0710	-.0214	.0674
2	.0414	-.7096	-.6593	-.2002	.1415

ZUORDNUNGSVERFAHREN

ANZAHL DER GRUPPEN 1
ANZAHL DER VARIABLEN 2
ANZAHL DER OBJEKTE 600
KREUZPRODUKTMATRIX W (INVERHALB DER GRUPPEN)

W-M SECTION 1

ROW 1 2
1 129.9823 -22.1068
2 -22.1068 295.9092

MATRIX DER GRUPPENMITTELWERTE

X-M SECTION 1

ROW 1 2 3
1 2.3882 .2210 .0969
2 .0908 .6509 .3186

T*2-TEST AUF UNTERSCHIEDE ZWISCHEN JE ZWEI GRUPPEN

FREIHEITSGRADE F1 = 2 F2 = 446

MATRIX DER D*2-WERTE (UNTERHALB DER DIAGONALEN)
MATRIX DER F-WERTE (UEBER DER DIAGONALEN)

F-M SECTION 1

ROW 1 2 3
1 1.0000 606.4744 677.0330
2 16.2089 1.0000 9.1284
3 48.0947 .2440 1.0000

MATRIX DER KOFFIZIENTEN DER ZUORDNUNGSFUNKTIONEN
(DIE LETZTE ZEILE ENTHAELT DIE KONSTANTEN)

C-M SECTION 1

ROW 1 2 3
1 .0166 .0177 .0012
2 -.0007 .0006 .0012
3 -.0214 -.0221 -.0008

ZUORDNUNG ZUR GRUPPE GR.NEU BEI GEGEBENER GRUPPENEINTEILUNG GR.ALT

OBJEKT	B1001	B1002	B1003	B1004	B1005	B1006	B1007	B1008	B1009	B1010	B1011	B1012	B1013	B1014	B1015
GR.ALT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GR.NEU	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
OBJEKT	B1016	B1017	B1018	B1019	B1020	B1021	B1022	B1023	B1024	B1025	B1026	B1027	B1028	B1029	B1030

GR.ALT GR.NEU	81031 4 4	81032 4 4	81033 1 1	81034 1 1	81035 1 1	81036 1 1	81037 1 1	81038 1 1	81039 1 1	81040 1 1	81041 1 1	81042 1 1	81043 1 1	81044 1 1	81045 1 1
OBJEKT GR.ALT GR.NEU	81046 1 1	81047 1 1	81048 1 1	81049 1 1	81050 1 1	81051 1 1	81052 1 1	81053 1 1	81054 1 1	81055 1 1	81056 1 1	81057 1 1	81058 1 1	81059 1 1	81060 1 1
OBJEKT GR.ALT GR.NEU	81061 1 1	81062 1 1	81063 1 1	81064 1 1	81065 1 1	81066 1 1	81067 1 1	81068 1 1	81069 1 1	81070 1 1	81071 1 1	81072 1 1	81073 1 1	81074 1 1	81075 1 1
OBJEKT GR.ALT GR.NEU	81076 1 1	81077 1 1	81078 1 1	81079 1 1	81080 1 1	81081 1 1	81082 1 1	81083 1 1	81084 1 1	81085 1 1	81086 1 1	81087 1 1	81088 1 1	81089 1 1	81090 1 1
OBJEKT GR.ALT GR.NEU	81091 1 1	81092 1 1	81093 1 1	81094 1 1	81095 1 1	81096 1 1	81097 1 1	81098 1 1	81099 1 1	81100 1 1	81101 1 1	81102 1 1	81103 1 1	81104 1 1	81105 1 1
OBJEKT GR.ALT GR.NEU	81106 1 1	81107 1 1	81108 1 1	81109 1 1	81110 1 1	81111 1 1	81112 1 1	81113 1 1	81114 1 1	81115 1 1	81116 1 1	81117 1 1	81118 1 1	81119 1 1	81120 1 1
OBJEKT GR.ALT GR.NEU	81121 1 1	81122 1 1	81123 1 1	81124 1 1	81125 1 1	81126 1 1	81127 1 1	81128 1 1	81129 1 1	81130 1 1	81131 1 1	81132 1 1	81133 1 1	81134 1 1	81135 1 1
OBJEKT GR.ALT GR.NEU	81136 1 1	81137 1 1	81138 1 1	81139 1 1	81140 1 1	81141 1 1	81142 1 1	81143 1 1	81144 1 1	81145 1 1	81146 1 1	81147 1 1	81148 1 1	81149 1 1	81150 1 1
OBJEKT GR.ALT GR.NEU	82001 2 2	82002 2 2	82003 2 3	82004 2 3	82005 2 2	82006 2 2	82007 2 3	82008 2 2	82009 2 2	82010 2 3	82011 2 2	82012 2 2	82013 2 3	82014 2 2	82015 2 2
OBJEKT GR.ALT GR.NEU	82016 2 3	82017 2 2	82018 2 3	82019 2 2	82020 2 2	82021 2 2	82022 2 3	82023 2 2	82024 2 2	82025 2 2	82026 2 3	82027 2 2	82028 2 2	82029 2 3	82030 2 2
OBJEKT GR.ALT GR.NEU	82031 2 3	82032 2 3	82033 2 2	82034 2 3	82035 2 3	82036 2 3	82037 2 2	82038 2 2	82039 2 3	82040 2 2	82041 2 2	82042 2 2	82043 2 3	82044 2 2	82045 2 3
OBJEKT GR.ALT GR.NEU	82046 2 2	82047 2 2	82048 2 2	82049 2 2	82050 2 3	82051 2 2	82052 2 2	82053 2 2	82054 2 3	82055 2 2	82056 2 3	82057 2 2	82058 2 3	82059 2 3	82060 2 2
OBJEKT GR.ALT GR.NEU	82061 2 2	82062 2 2	82063 2 2	82064 2 2	82065 2 3	82066 2 2	82067 2 3	82068 2 2	82069 2 3	82070 2 3	82071 2 2	82072 2 2	82073 2 3	82074 2 3	82075 2 2
OBJEKT GR.ALT GR.NEU	82076 2 2	82077 2 2	82078 2 2	82079 2 3	82080 2 3	82081 2 3	82082 2 2	82083 2 2	82084 2 3	82085 2 3	82086 2 2	82087 2 2	82088 2 2	82089 2 2	82090 2 3
OBJEKT GR.ALT GR.NEU	82091 2 2	82092 2 2	82093 2 2	82094 2 3	82095 2 3	82096 2 2	82097 2 2	82098 2 2	82099 2 2	82100 2 3	82101 2 2	82102 2 2	82103 2 2	82104 2 2	82105 2 2

OBJEKT	81196	81197	81198	81199	81200	82151	82152	82153	82154	82155	82156	82157	82158	82159	82160
GR.ALT	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
GR.NEU	1	1	1	1	1	2	3	3	3	3	3	3	2	3	2
OBJEKT	82161	82162	82163	82164	82165	82166	82167	82168	82169	82170	82171	82172	82173	82174	82175
GR.ALT	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
GR.NEU	3	2	3	3	3	2	2	2	3	3	3	2	2	3	2
OBJEKT	82176	82177	82178	82179	82180	82181	82182	82183	82184	82185	82186	82187	82188	82189	82190
GR.ALT	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
GR.NEU	3	3	3	3	3	3	2	2	2	3	3	2	3	2	3
OBJEKT	82191	82192	82193	82194	82195	82196	82197	82198	82199	82200	83151	83152	83153	83154	83155
GR.ALT	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
GR.NEU	2	3	2	2	3	3	3	2	2	2	3	2	3	2	3
OBJEKT	83156	83157	83158	83159	83160	83161	83162	83163	83164	83165	83166	83167	83168	83169	83170
GR.ALT	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
GR.NEU	1	3	2	3	2	3	3	3	2	3	3	3	3	2	3
OBJEKT	83171	83172	83173	83174	83175	83176	83177	83178	83179	83180	83181	83182	83183	83184	83185
GR.ALT	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
GR.NEU	3	3	2	3	1	3	3	3	3	2	2	3	3	1	3
OBJEKT	83186	83187	83188	83189	83190	83191	83192	83193	83194	83195	83196	83197	83198	83199	83200
GR.ALT	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
GR.NEU	1	3	1	3	3	2	3	3	3	3	3	3	2	2	2

0 56 60 150
 FEHLER
 18 127

Danksagung

Im folgenden möchte ich mich bei meinen Doktorvätern Prof.Dr.-Ing. D.Popović und Prof.Dr.S.Boseck für die Betreuung und Unterstützung während der Promotionszeit bedanken.

Für nützliche Bemerkungen vom mathematischen Gesichtspunkt, die zur Endfassung dieser Arbeit geführt haben, richte ich meinen besonderen Dank an Herrn Prof.Dr.J.Timm.

Die mit meinen Kollegen, Dipl.Ing.P.Anders, Dipl.Ing. W.Bornemann und Dipl.Ing.K.Harbers, durchgeführten fachlichen Diskussionen haben bei der Koordinierung der Forschungsarbeit und beim Zusammenschreiben der Dissertation eine wichtige Rolle gespielt. Hierbei möchte ich mich für viele nützliche Hinweise bei Herrn Dr.Ing.U.Christ bedanken.

Für die bei der Erstellung dieser Arbeit notwendigen Schreibarbeiten gilt mein besonderer Dank Frau A.Andrei.

